

**Министерство образования и науки  
Украины  
Харьковский национальный университет  
им. В.Н. Каразина**

**А.П. Козлов, Н.Н. Попов**

# **МЕДИЦИНСКАЯ СТАТИСТИКА**

Учебное пособие

**ХАРЬКОВ  
2006**

## **УДК 612.017.1**

**Козлов А.П., Попов Н.Н. Медицинская статистика: учебное пособие.** – Харьков, издат. центр ХНУ, 2006. – 88 с.

В учебном пособии приводятся основные положения медицинской статистики, основанные на программе курса социальная медицина, экономика и организация здравоохранения. В пособии приведены современные методы статистической обработки медико-биологических данных, которые могут быть использованы для выполнения курсовых работ, научных работ, диссертаций. Для каждого из методов медико-статистической обработки данных приведено и обосновано их применение в практическом здравоохранении. В конце каждой из приведенных глав приведены решения эталонных задач, даны задачи для самостоятельной работы студентов. Для студентов, интернов и врачей всех специальностей.

### **Рецензенты:**

Волкославская В.М. – профессор, доктор медицинских наук, заведующая отделом научно-аналитической работы института дерматологии и венерологии

Просоленко Н.В. – кандидат медицинских наук, доцент кафедры социальной медицины, экономики и организации здравоохранения ХМАПО

**Рекомендовано к печати Ученым советом факультета фундаментальной медицины ХНУ им. В.Н. Каразина, протокол №8 от 19 апреля 2006 года.**

® Козлов А.П., Попов Н.Н.  
© Козлов А.П., Попов Н.Н.  
© ХНУ им. В.Н. Каразина,  
2006

## **Предисловие**

Полная и достоверная статистическая информация является тем необходимым основанием, на котором базируется процесс управления и организация здравоохранения во всех странах мира. Вся информация, имеющая медицинскую значимость, в конечном счете, обрабатывается и анализируется с помощью медицинской статистики.

Именно медико-статистические данные позволяют определить заболеваемость по отдельным видам патологии, выделить наиболее приоритетные для профилактической работы врачей направления, выявить основные тенденции развития заболеваний, прогнозировать развитие эпидемий, оценить уровень риска у больного с той или иной патологией, проанализировать состояние больного и выбрать наиболее рациональный метод лечения, исследовать уровень здоровья, как отдельного пациента, так и выбранной группы (вплоть до здоровья нации), исследовать уровень жизни населения и его взаимосвязь с уровнем здоровья, определить влияние факторов внешней среды на состояние организма пациента и другие медико-социальные, медико-биологические, медико-экономические и медико-статистические явления и процессы.

Овладение статистической методологией - одно из условий познания конъюнктуры медицинского рынка, изучения тенденций и прогнозирования, принятия оптимальных решений на всех уровнях деятельности.

Сложной, трудоемкой и ответственной является заключительная, аналитическая стадия исследования. На этой стадии рассчитываются средние показатели и показатели распределения, анализируется структура совокупности, исследуется динамика и взаимосвязь между изучаемыми явлениями и процессами.

На всех этапах статистического исследования статистика использует различные методы. Поэтому полное познание основ медицинской статистики является неотъемлемым условием успешной медицинской практики.

## **Социальная медицина и организация здравоохранения. Медицинская статистика. Организация статистического исследования**

Во время перемен, которые происходят в нашем обществе, острой необходимостью становится разработка и внедрение оптимальной системы здравоохранения, и управление ею с помощью современных методов, использующихся в мировой теории и практике.

Наша наука и соответствующая ей учебная дисциплина до настоящего времени, пожалуй, вызывает много вопросов, что это такое. Начнем с названия, которое до сих пор дискуссионно.

Существует несколько предназначенных для названия социальной медицины терминов: *общественная медицина и организация здравоохранения (Россия), социальная гигиена и организация здравоохранения (страны Восточной Европы), в англоязычных странах этот предмет чаще называют общественным здоровьем, во франкоязычных – социальной медициной, в США ранее, чем в других странах, его стали обозначать как социология медицины или социология здравоохранения.*

Названия большей части научно-исследовательских учреждений по нашей специальности связаны именно с понятием социология. Однако в «чистом виде», кроме ряда лабораторий и кафедр в США, предмет представлен нечасто. Обычно он сочетается с другими дисциплинами или разделами науки: медицинской статистикой, эпидемиологией, особенно неэпидемических заболеваний, гигиеной, историей медицины и здравоохранения, медицинское право, социальное страхование, экономика и даже тропические болезни. Так, крупный исследовательский центр по социальной медицине традиционно существует в Лондонском институте тропической медицины. Специальные НИИ по проблемам здравоохранения обозначаются институтами гигиены и здравоохранения – в Софии и Бухаресте, эпидемических

исследований во Франции, национальным центром статистики в США.

В настоящее время преподавание дисциплины осуществляется в абсолютном большинстве медицинских учебных заведений мира, хотя, как и следовало ожидать, не всегда как самостоятельный предмет. Самостоятельные кафедры организованы практически во всех странах и содержат, как правило, курсы социологии, статистики здравоохранения, эпидемиологии неэпидемических болезней, информатики, организации работы медицинских учреждений, управления (менеджмента), страхования здоровья и др. С 20-х годов, со времени создания первых кафедр и курсов, создаются учебники, руководства и другие пособия по предмету. Они выходят чаще всего под названиями «Социальная медицина», «Медицинская социология», «Социология медицины».

В отличие от большинства клинических дисциплин, имеющих дело с отдельным человеком, индивидуумом и его здоровьем, предмет социальная медицина изучает состояния и охрану здоровья населения, то есть он непосредственно сталкивается с общественными социальными проблемами и процессами, т.о. занимая место «моста» между медициной и общественными дисциплинами, прежде всего с социологией. Социальная медицина по существу является наукой о стратегии и тактике здравоохранения.

### **Социальная медицина и организация здравоохранения**

– это наука и учебная дисциплина, изучающая закономерности общественного здоровья и систему его охраны.

**Целью** изучения является разработка и научное обоснование социальных, медицинских и экономических мероприятий по обеспечению оптимального уровня здоровья населения и его активного долголетия.

### **Задачи:**

1. Изучение состояния здоровья населения и процессов его воспроизводства;
2. Обеспечение характеристики динамики здоровья населения, как в целом, так и отдельных его групп;

3. Выявление условий и факторов, которые приводят к положительным и отрицательным отклонениям в состоянии здоровья различных групп населения;
4. Разработка направлений оздоровления населения и определение принципов системы охраны здоровья, его теоретических и организационных основ;
5. Анализ деятельности органов и учреждений здравоохранения, создание их рациональных структур и научное обоснование наиболее приемлемых форм организации работы, проведение реформирования и реконструкции;
6. Создание прогнозов и планов развития системы здравоохранения с целью проведения целенаправленных мероприятий по поддержанию должного уровня состояния здоровья населения.

**Социальная медицина и организация здравоохранения состоит из следующих разделов:**

- Медицинская статистика;
- Здоровье населения и факторы окружающей среды, влияющие на него;
- Система здравоохранения.

Наиболее характерной особенностью методологической базы является ее широта и многообразие методик, взятых из других дисциплин. Таким образом, необходимо говорить не о специфике научных методов, а о своеобразии объектов исследований (больные, учреждения здравоохранения, медицинский персонал и др.

**Методы социальной медицины:**

- *Статистический* – основной метод социальной медицины т.к. он позволяет изучить массовые явления, касающиеся здоровья и его охраны;
- *Социологический* – позволяет изучить социальную структуру общества и ее влияние на здоровье, методика базируется на анкетировании и интервьюировании;
- *Эпидемиологический* – базируется на изучении многолетней динамики заболеваемости по отдельным видам патологии и заболеваний, изучении правил и законов течения эпидемий;
- *Экспертных оценок* – с его помощью изучается качество и эффективность медицинской помощи;

- *Исторический* – устанавливает исторические закономерности развития общественного здоровья и его охраны;
- *Экономический* – дает возможность оценить экономическую эффективность систем медицинской помощи;
- *Экспериментальный* – позволяет изучать преимущества и недостатки организационных форм медицинской помощи;
- *Системного анализа* - используется при исследовании общественных и биологических систем (совокупности взаимосвязанных частей, элементов).
- *Изучение общественного здоровья* - представляет собой сложную, развивающуюся, динамичную систему, тесно связанную с другими системами, отраслями, областями знаний.

### **История возникновения и развития предмета.**

Предпосылкой возникновения науки и учебной дисциплины об общественном здоровье и здравоохранении было возникновение общественной потребности научно обоснованного объяснения природы здоровья и болезней населения с целью направленного, регулируемого воздействия, т.е. управления им. Таким образом, формированию социальной медицины как науки предшествовал длительный период накопления знаний о мероприятиях по охране и укреплению здоровья общества.

Еще в 17 веке англичане Джон Граунт (1620-1674) и Вильям Петти (1623-1687) внедрили статистику в медицину (в изучение демографии).

В 18 веке на медицинских факультетах Западной Европы начинают впервые включать как отдельный раздел в состав той или иной медицинской дисциплины, чаще всего в судебную медицины, так называемую медицинскую полицию, которая освещала санитарные мероприятия общественного характера. Прежде всего, эпидемических заболеваний. Медицинская полиция стала первым обобщением представлений об обязанностях государства по охране здоровья своих граждан, главным образом в виде полицейских актов и запретов. Идеи медицинской полиции развивались в России многими учеными, в

частности Е. Мухиным, П. Пелехиным, Ф. Керестури, который начал систематическое преподавание медицинской полиции в России. Последнюю, дополненную медико-топографическими и статистическими описаниями, следует рассматривать как ближайшую предшественницу социальной медицины.

*Основателем социальной медицины в мире* считается австрийский врач клиницист и гигиенист *Йоганн Петер Франк* (1745-1821), который в своем труде «Система совершенной медицинской полиции» отделил вопросы общественной охраны здоровья, провозгласил потребность рационального законодательства для борьбы с массовыми заболеваниями, организации лечения и родовспоможения, пропаганды гигиенических навыков среди населения. И.П.Франк применил статистику для доказательства важности общественной охраны здоровья, и также разработал модель системы медицинской опеки «от колыски до могилы». То есть он впервые обосновал и научно разработал идею государственной организации медицинской помощи. И.П.Франк внес большой вклад в развитие медицинской и общественной науки в России: в 1805-1808 гг. он был ректором Петербургской медико-хирургической академии, где развивал свои идеи общей медицинской полиции.

Следует отметить важную роль И.А.Данилевского, который в 1784 г. В Геттингене защитил диссертацию «О государственной власти как опытнейшем враче». В ней Данилевский широко освещает вопросы о задачах государства в деле охраны здоровья населения, особенно подчеркивает важность распространения санитарного просвещения среди населения с детства, необходимость ввести во всех школах преподавание «первооснов сохранения здоровья». Свою диссертацию он заканчивает таким выводом: «Излечение заболеваний, искоренение их причин необходимо ожидать не от врачей и аптекарей, а только от государственной власти».

Особые заслуги в борьбе за внедрение в жизнь оздоровительных мероприятий государственного характера принадлежат англичанам С.Смиту (1788-1861) и Е.Чедвику (1800-1890). Под их влиянием в Англии было



создано Генеральное управление здоровья (1848), где начался научный анализ статистических ведомостей, впервые внедрены должности санитарных врачей.

Однако до начала 20 века не существовало специальных курсов, кафедр социальной медицины, не было периодических изданий. Сначала научные и учебные учреждения этого профиля возникли в Германии, где уже в начале 70-х годов 18 века были внедрены некоторые реформы в здравоохранении (внедрено социальное страхование за счет средств предприятий, государственного бюджета и трудящихся, организованы амбулатории, некоторые диспансеры и др.). С 1903 года тут издается первый журнал по социальной гигиене. В 1905 основано в Берлине Научное общество социальной гигиены и медицинской статистики. Это общество способствовало созданию организаций, которые занимались вопросами охраны здоровья детей, борьбы с туберкулезом и алкоголизмом в социально-медицинском аспекте, однако оно просуществовало до начала 1 мировой войны.

Научные основы социальной медицины были заложены немецким ученым Алфредом Гротьяном (1869-1921), который в 1920 году организовал и возглавил в Берлине первую самостоятельную кафедру. В своем капитальном труде «Социальная патология» А.Гротьян выделяет основные группы заболеваний с точки зрения их социальной обусловленности и обозначает пути общественной борьбы с ними. Позже и в других университетах Германии стали создаваться подобные курсы и кафедры, на которых работали известные социал - гигиенисты Э.Ресле, Ф.Принцинг, А.Фишер.

Впервые понятие «социальная медицина» ввел французский ученый J.Guerin (1848). Под этим он понимал изучение связи физического и психического состояния населения и законодательства, а также других социальных институтов, изучение связи между социальными факторами, здоровьем и заболеваемостью, выработку мер по укреплению здоровья и профилактики заболеваний.

В 1864 году Gasper были опубликованы данные о средней продолжительности жизни немецкого дворянства и берлинской городской бедноты. Аналогичные данные опубликовали в тот же период в Англии Cley (1844), Gavin (1848), Farr (1864). В Германии в прошлом веке работали выдающиеся деятели в области социальной медицины – R.Virchow, S.Neman. Медицина была объявлена социальной наукой, а работа Вирхова «Сообщения об эпидемии тифа в Верхней Силезии» является классическим трудом по немецкой социальной гигиене. Неман в работе «Общественная забота о здоровье и собственность» заявляет, что медицина относится к социальным наукам.

Одним из значительных произведений мировой статистики первой половины 20 века был труд выдающегося украинского экономиста, теоретика социальной статистики Д.Журавского «Об источниках и использовании статистических ведомостей» (1846). Разработанная автором теория сбора, обработки и анализа статистических материалов базировалась не только на количественных, но и на качественных характеристиках общественных явлений в их причинной обусловленности с учетом отдельных социальных групп и классов.

В 18-19 веках были проведены медико-топографические исследования санитарного состояния многих отдельных городов и губерний. Они были первичными формами статистических исследований. Новой ступенью развития описаний стали демографические и санитарно-статистические работы в период земской медицины, тогда же сформировалось социально-медицинское направление в общественной медицине. Среди организаторов санитарного дела в Украине того времени следует вспомнить М.И.Тезякова (1859-1925) и С.М.Игумнова (1864-1942), М.Уварова, П.Кудрявцева, П.Диатропова. Работая санитарными врачами уездов, они изучали влияние условий труда и быта на здоровье сельскохозяйственных работников, придавая большое значение профилактическим мероприятиям на врачебных участках. Земские врачи Украины внесли большой вклад в разработку новых форм медико-санитарного обеспечения населения. Медицинская комиссия Полтавского губернского земства в 1869 г. сформулировала задачи

здравоохранения, которые перекликаются с организационными формами сельского здравоохранения в более поздний период. Херсонские врачи организовывали медико-продовольственные пункты, пункты регистрации пришлых сельскохозяйственных работников, их примеру последовали другие земства.

*В 1874 г. 1 съезд земских врачей Херсонской области сформулировал основные принципы организации земской медицины:*

- Доступность медицинской помощи для всех слоев населения;
- Бесплатность медицинской помощи для пациентов;
- Квалифицированность медицинской помощи;
- Санитарный (профилактический) характер;
- Участковый принцип организации с разворачиванием на каждом участке амбулатории и стационара;
- Регистрация болезней и смертей с целью изучения их распространенности.

С.Игумнову принадлежит идея создания санитарных станций – принципиально новой организационной формы санитарно-противоэпидемической работы. Харьковское земство организовывало передвижные гигиенические выставки с целью распространения гигиенических знаний среди народа.

Формирование самостоятельных гигиенических дисциплин в Украине, их общественная пропаганда тесно связаны с деятельностью Киевского и Харьковского университетов, в которых во второй половине 19 века были созданы первые кафедры этого направления (И.П.Скворцов, В.А.Субботин). Преподавание социально-медицинских вопросов и статистики началось с создания в 1871 г. Первой в Украине кафедры гигиены, медицинской полиции, медицинской географии и статистики при Киевском университете. В 1907 году О.В.Корчак-Чепурковский (1857-1947). читал курс «Основы социальной гигиены и общественной медицины», составил одну из первых программ по этой дисциплине. Следует подчеркнуть, что как в Киеве, так и в Петербурге и Москве преподавание таких курсов было начато в учебных учреждениях немедицинского профиля (Коммерческом институте,

политехническом и т.д.). Показательной для периода командно-административной системы была судьба О.В.Корчака-Чепурковского, который возглавлял кафедры гигиены в медицинских учебных заведениях Киева. В 1921 году он первым из медиков был избран действительным членом Всеукраинской академии наук по специальности социальная медицина (после него по этой специальности в Украине 80 лет больше никто не избирался). В середине 30-х годов подвергся гонениям со стороны правительства за публикацию научного труда о смертности населения Киева, Харькова, Одессы и Днепропетровска, который отражал последствия искусственного голодомора в Украине. После этого ему разрешили работать только в отделе демографии и санитарной статистики, его научные труды не публиковались. А в 1936 г. Был обвинен в фашистско-националистических взглядах его сын, обвинителем которого стал проф. С.С.Каган, ученик его отца. Через некоторое время по доносу арестован и сам С.С.Каган.

В Украине первые самостоятельные кафедры появились во второй половине 1923 г. В 1922 г. Было предложено разделить преподавание гигиены на три части: общую гигиену, социальную и профессиональную. В содержание социальной медицины предлагалось включить историю ее развития, демографию с медицинской статистикой, учение о социальных болезнях, общественную охрану здоровья в России и на Западе.

Первая самостоятельная кафедра социальной гигиены создана в Харьковском медицинском институте (зав. каф. М.Гуревич), тогда же аналогичная кафедра была создана в Одесском медицинском институте (зав.каф. Л.Громашевский) и Киевском медицинском институте (зав. каф.С.С.Каган). В 1924 г. кафедра была организована в Днепропетровском медицинском институте (зав. каф. М.Доныч). Своеобразие преподавания социальной гигиены в Украине заключалось в том, что социально-гигиенические вопросы, кроме основного курса, входили в курсы социальной борьбы с туберкулезом, венерическими заболеваниями, которые преподавались при соответствующих клиниках, в программы летнего практикума, стажировки после окончания ВУЗа и в

тематику дипломных работ, которые защищались после годичной стажировки.

С 1923 г. начинают проводиться массовые санитарно-демографические обследования для получения научных данных и разработки на их основе практических оздоровительных мероприятий (размеры эпидемий тифов, изучение сельских домов, водоснабжения, условий труда, питания, санитарной грамотности). Исследования дали материал для серьезных научных обобщений и с этой целью начали в 23-25 гг. создаваться научно-исследовательские институты с отделами социальной патологии и гигиены. В 1920-30 гг. создались два центра развития социальной медицины как науки и предмета преподавания: Харьковский во главе с С.Томилиным (М.Гуревич, З.Гуревич, И.Арнольди, А.Мерков, П.Петров, С.Екель и др.) и Киевский во главе с С.Каганом (Б.Шкляр, И.Овсиенко, К.Дупленко, Е.Белицкая, Л.Лекарев.) В работах ученых большое внимание уделялось медицинской статистике, исследованию медико-демографических процессов и влиянию на них условий жизни, истории медицины. Были подготовлены первые украинские учебники по социальной гигиене: «Очерки теории социальной гигиены» С.Кагана (1932 г.), «Социальная гигиена» С.Томила и А.Меркова (1933).

В 1934 г. В украинском демографическом институте был создан отдел санитарной статистики, учреждение переименовано в Институт демографии и медицинской статистики. До начала 30-х годов был накоплен большой опыт разработки вопросов теории и практики социальной гигиены. Но с начала 30-х гг. с усилением административно-командной системы социальная гигиена стала предметом жесткой критики, вследствие чего постепенно социально-гигиенические исследования фактически прекратились. В 1934 г. Ликвидировано 38 НИИ, в 1938 г. закрыт Институт демографии и медицинской статистики. В мае 1941 г. кафедры были переименованы в кафедры организации здравоохранения, что сузило круг их задач. На 25 лет было запрещено проведение социально-медицинских исследований и преподавание предмета.

В 1966 г. социальная гигиена была восстановлена как наука и предмет преподавания, кафедры переименованы в кафедры социальной гигиены и организации здравоохранения. Однако, только в 1989 г. В Украине была восстановлена работа единой научно-исследовательской институции – Института социальной гигиены и управления здравоохранением в Киеве. С 1996 г. Этот институт носит название «Институт общественного здоровья». В 1990 г. 2 съезд социальных гигиенистов Украины сменил название науки на социальная медицина и организация здравоохранения.

Современные зарубежные представители нашей науки, такие как Р.Занд, У.Уинслоу, А.Паризо, Л.Поппер, К.Канапериа, а также создатели новейших школ – Р.Дюбо, К.Эванг, П.Делор, Х.Дон, Т.Персон, Э.Фрейдсон, Д.Механик, Л.Бернар, М.Кандау, Х.Мадлер и др. продолжают линию на выявление социальной обусловленности общественного здоровья, стоят на позициях реформ охраны здоровья населения, примат государственной, правительственной медицины. Один из лидеров социальной медицины Рене Занд рассматривал социальную медицину как ключ к погашению конфликтов, характерных для современного общества.

В XX веке социальная медицина и организация здравоохранения выходит за рамки национальных границ и становится международной.

Проблемы здоровья и его охраны разрабатываются многими международными организациями. Ведущую роль, среди которых играет ВОЗ. Эта организация является одним из подразделений ООН. Решение о ее создании принято в 1946 г., а устав ратифицирован 7 апреля 1948 г. И с тех пор этот день отмечают как Всемирный день здоровья.

Главным положением устава ВОЗ является достижение всеми народами наивысшего возможного уровня здоровья.

### **Основные задачи ВОЗ:**

- Координация международной деятельности в области охраны здоровья;

- Предоставление странам необходимой информации о здоровье людей и состоянии его охраны;
- Предоставление странам помощи в вопросах охраны здоровья;
- Организация работы по борьбе с эпидемическими, эндемическими болезнями другими болезнями;
- Охрана психического здоровья;
- Проведение систематических исследований в области охраны здоровья;
- Способствование подготовке медицинских кадров.

Высшим органом ВОЗ является Всемирная ассамблея охраны здоровья, которая собирается в Женеве ежегодно. Между сессиями ассамблей работает исполнительный комитет, состоящий из представителей 30 государств. Текущую организационную работу проводит секретариат ВОЗ, возглавляемый Генеральным директором.

В составе ВОЗ 6 региональных бюро: африканское (центр в Браззавиле), американское (Вашингтон), Северо-Восточной Азии (Дели), Европы (Копенгаген), Восточного Средиземноморья (Александрия), Западной части Тихого океана (Манила).

Ежегодный бюджет ВОЗ достигает более 1 миллиарда долларов США. Эти средства выделяют государства-члены ВОЗ и пополняются различными гуманитарными и добровольными взносами.

Основная деятельность ВОЗ реализуется через исполнение различных программ на международном и национальном уровнях. Самыми известными программами стали программа борьбы с малярией, программа ликвидации натуральной оспы, внедрения системы первичной медико-санитарной помощи, подготовки медицинских кадров.

Украина принимает участие в работе ВОЗ, имея право голоса. Сотрудничество на протяжении последних лет укрепилось, несмотря на то, что с 1992 г. Украина фактически является должником ВОЗ.

Наша страна принимает участие в программе ВОЗ «Евроздоровье», которая началась в 1991 г. И направлена на активное сотрудничество и помощь странам Восточной,

Центральной и Южной Европы. Так по краткосрочной программе Украине была оказана помощь по борьбе с холерой. По средне- и долгосрочным программам предусмотрена помощь в реформировании системы здравоохранения, ликвидации полиомиелита.

Ежегодная помощь ВОЗ нашей стране составляет 50 тыс. долларов.

Для координации международной деятельности в области здравоохранения ВОЗ в начале 80-х разработала программу «Стратегия достижения здоровья для всех до 2000 г.», которая с 1997 г. Получила название «Достижение здоровья для всех в 21 веке». *Основными целями Европейской программы являются:*

- Достижение справедливости и равных прав по вопросам охраны здоровья для населения разных Европейских стран;
- Укрепление здоровья наиболее уязвимых групп населения (инвалиды, лица пожилого возраста, дети, молодежь, женщины, лица с хроническими заболеваниями);
- Значительное уменьшение наиболее распространенных заболеваний (сердечно-сосудистых, злокачественных новообразований, травматизма, психических, инфекционных).

*Пути устранения разницы в состоянии здоровья:*

- Внедрения мониторинга различий в показателях различных стран и социальных групп;
- Обеспечение большей доступности социальных условий: надлежащее питание, жилищные условия, образование, медико-санитарная помощь;
- Больше внимание странам, находящимся в неблагоприятных условиях.

Как мы видим, многими поколениями ученых формировалась современная точка зрения, что человек является **социопсихобиосистемой**, здоровье которой во многом зависит от социальных факторов.

**Статистика** – самостоятельная наука, которая изучает количественную сторону общественных явлений в неразрывной связи с их качественной стороной в



конкретных исторических условиях места и времени. Основа статистики – теория вероятности.

Отрасль статистики, которая изучает вопросы, связанные с медициной, гигиеной и общественной охраной здоровья, называют **медицинской статистикой**.

Выделяют **3 раздела медицинской статистики**:

- статистика здоровья населения
- статистика здравоохранения
- статистическая оценка медико-биологических данных

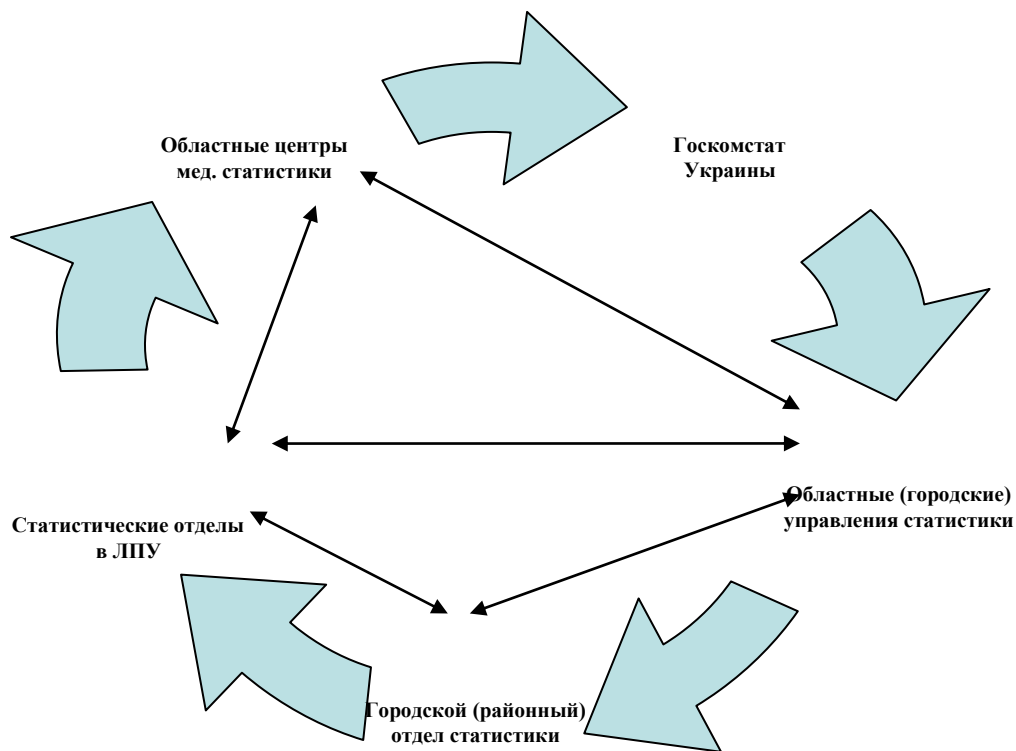
**Задачи медицинской статистики:**

- Изучение здоровья населения, как в целом, так и отдельных его групп.
- Выявление и установление связей уровня общей заболеваемости и смертности населения с различными факторами окружающей среды.
- Изучение данных о кадрах, сети медицинских учреждений.
- Оценка эффективности деятельности ЛПУ и лечебно-диагностического процесса.
- Установление достоверности результатов выборочных статистических исследований.

**Теоретическими основами медицинской статистики служат:**

- диалектика (все события рассматриваются во взаимосвязи, развитии, переходе из количества в качество)
- экономическая наука (оценка экономической эффективности мероприятий)
- медицинская наука
- общая теория статистики

Украина перешла на международную систему учета и статистики, внедряется мониторинг многих актуальных проблем, создана единая структурная информационная система учета состояния здоровья граждан Украины. Схематично она выглядит так:



Обмен информацией, сбор статистической информации и реализация статистического исследования требует **источника статистической информации**. Таковыми в большом количестве случаев в медицинском статистическом исследовании являются **учетные документы и другие источники получения информации для медико-статистического анализа**:

- статистический талон
- карта выбывшего из стационара
- экстренное сообщение о случае острого инфекционного заболевания
- паспорт ЛПУ
- акты, протоколы и т.д.

Изучение явлений с применением статистического метода требует умелого подхода к объекту исследования. **Требования к статистической информации** таковы, что она должна быть:

- однородной
- достоверной
- полной
- своевременной

**Объект статистического исследования** – статистическая совокупность. **Статистическая совокупность** – это группа, состоящая из большого количества относительно однородных элементов (**единиц наблюдения**), взятых вместе в конкретных границах пространства и времени.

Каждый элемент, составляющий статистическую совокупность и наделенный признаками сходства, называется **единицей наблюдения** и обозначается – п. Единица наблюдения – носитель учетных признаков, то есть признаков, по которым различаются элементы совокупности.

Признаки сходства служат основанием для формирования совокупности, признаки различия служат предметом анализа.

**По характеру признаки делятся на:**

- атрибутивные (описательного характера) – пол, наличие заболевания, диагноз, степень тяжести и т.д.
- количественные – уровень АД, температура тела и т.д.

Каждая величина количественного признака называется **вариантой** и обозначается V.

Исследователи изучают роль каждого признака во взаимосвязях изучаемого явления.

Следует различать факторные и результативные признаки. **Факторные признаки** – признаки, под влиянием которых изменяются другие признаки (**результативные**). Например, возраст – рост, метод лечения – результат. В каждой совокупности учетные признаки могут принимать различные значения.

Статистическая совокупность может быть генеральной и выборочной. **Генеральная совокупность** состоит из всех единиц, которые могут быть к ней отнесены. Однако, проведение такого исследования экономически невыгодно, требует значительных затрат времени, средств, хотя это наиболее вероятный метод.

**Выборочная совокупность** – это часть генеральной, отобранная специальным выборочным методом.

**Методы отбора выборочной совокупности** из генеральной:

- *случайный* – все единицы генеральной совокупности имеют равную вероятность попасть в выборку;
- *механический отбор* – по какому-нибудь признаку (алфавиту);
- *типологический* – выборка раскладывается на типичные группы (по полу, возрасту);
- *серийный* – вся выборка раскладывается на однородные группы (серии);
- *парно-сопряженный* – каждой единице наблюдения подбирается копия из контрольной группы;
- *метод направленного отбора* – выбирается исключительно однородная группа (по полу, возрасту, месту жительства);
- *когортный* – выбирается относительно однородная группа элементов, объединенных наступлением определенного признака;
- *анкетный* – информацию получают путем анализа заполненных анкет и отбора по определенному признаку;
- *anamnestический* – информация получена на основании опроса пациентов (единиц наблюдения) и выделения общей особенности.

**Свойства выборочной совокупности:**

- репрезентативность относительно генеральной
- достаточность по объему

*Репрезентативность* достигается правильным отбором единиц наблюдения, то есть, чтобы признаки были характерны для генеральной совокупности.

Объем выборки рассчитывается по формулам.

*Теоретическое обоснование этого метода* - теория вероятностей и закон больших чисел.

*Вероятность* – мера возможности возникновения какого-либо события в данных конкретных условиях, обозначается –  $p$ , изменяется в пределах от 0 до 1.

$p = m/n,$
------------

где  $m$  – наступление события,  $n$  – число всех возможных случаев.

Теория вероятностей обосновывает закон больших чисел: по мере увеличения числа наблюдений результат исследования стремиться воспроизвести данные генеральной совокупности; при достижении определенного числа наблюдений в выборочной совокупности результаты исследования будут максимально приближаться к данным генеральной совокупности.

**Статистическая совокупность обладает следующими свойствами:**

- распределение признака
- средний уровень признака
- разнообразие признака
- репрезентативность признака
- взаимосвязь между признаками

**Статистическое исследование** – это научно-организационный процесс, в котором по единой программе проводится наблюдение за определенными явлениями, сбор и регистрация первичных данных, их обработка и анализ.

**Этапы статистического исследования:**

1. Определение цели, составление программы и плана исследования.
2. Сбор материала.
3. Разработка данных.
4. Анализ, выводы, предложения.

**1 этап** – это проект всей работы. Необходимо определить цель исследования. Программа указывает направление работы, а план организацию исследования.

*Программа* включает сведения, которые необходимо собрать. Выделяют программу сбора материала, программу разработки, программу анализа.

*Программа сбора материала* заключается в определении официального документа или специально разработанной карты (анкеты) для регистрации данных в процессе наблюдения.

Программа разработки материала заключается в группировке учитываемых признаков (типологической, вариационной), составлении макетов таблиц (простой, групповой, комбинационной). В таблице выделяют подлежащее (признак явления по горизонтали) и сказуемое (признаки характеризующие подлежащее). Таблицы делятся в зависимости от своей сложности на:

- **Простые** – таблица имеет 1 подлежащие и 1 сказуемое. Пример ниже.

Название заболевания	Количество по участку №1
Пневмония	23
Острый бронхит	56
Бронхиальная астма	11

- **Групповые** – таблица имеет 1 подлежащие и несколько сказуемых. Пример ниже.

Название заболевания	Количество по участку №1	Пол		Возраст	
		М	Ж	До 25 лет	Более 25 лет
Пневмония	23	20	3	14	9
Острый бронхит	56	34	22	50	6
Бронхиальная астма	11	6	5	1	10

- **Комбинационные** – таблица имеет подлежащие и одно или несколько взаимосвязанных сказуемых. Пример ниже.

Название заболевания	Возраст					
	0 – 33 лет		33 – 66 лет		66 – 99 лет	
	М	Ж	М	Ж	М	Ж
Пневмония	12	11	22	44	11	45
Острый бронхит	1	6	34	33	15	43
Бронхиальная астма	4	9	11	13	34	43

На основании таблиц составляют *программу анализа*.

Составление плана исследования: определение объекта наблюдения (население, работающие, школьники); способов формирования статистической совокупности (по охвату совокупность может быть сплошной или выборочной, по времени – текущей и единовременной; по виду наблюдения – непосредственное или выкопировка сведений); приемы разработки материала (ручной, машинный); место и сроки проведения работы; силы, средства и т.д.

**2 этап** – сбор материала – статистическое наблюдение, регистрация в соответствии с программой сбора материала.

**3 этап** – шифровка, раскладка по группам, подсчет карт, заполнение таблицы, вычисление статистических показателей, графическое изображение полученных данных.

**4 этап** – проведение анализа (интерпретация и сравнение данных), литературное оформление, оформление выводов и предложений для внедрения в практику.

### **Задача – эталон №1**

**Условие.** При анализе заболеваемости населения на участке за 1 год врач – терапевт составил несколько макетов статистических таблиц. Составьте групповую таблицу: «Распределение больных с различными нозологическими формами по полу и возрасту».

**Решение:**

В задании необходимо построить групповую таблицу, т.е. таблицу, которая имеет 1 подлежащие и несколько сказуемых. Подлежащим в данном случае является нозологическая форма, а сказуемыми – пол и возраст больных. Исходя из этих соображений, строим групповую таблицу:

**Таблица №1. «Распределение больных с различными нозологическими формами по полу и возрасту»**

Название заболевания	Пол		Возраст	
	М	Ж	До 25 лет	Более 25 лет
Пневмония	20	3	14	9
Острый бронхит	34	22	50	6
Бронхиальная астма	6	5	1	10

### **Задача – эталон №2**

**Условие.** Составить план медико-статистического исследования по разработке плана профилактических мероприятий на основании обращений к терапевту.

**Решение.** Согласно приведенного ранее плана медико-статистического исследования составим план данного исследования:

*Цель исследования:* Разработать мероприятия по

профилактике заболеваний на основании обращений к терапевту

*Задачи исследования:*

1. Изучить состав больных по полу и возрасту
2. Изучить структуру заболеваемости
3. Изучить структуру заболеваемости по отдельным нозологическим формам
4. Разработать мероприятия по снижению уровня и профилактике заболеваний

*Программа исследования:*

1. Единица наблюдения – пациент на приеме у терапевта
2. Изучаемые признаки – нозологическая форма, пол, возраст, цель посещения, случай обслуживания
3. Учетный документ – талон амбулаторного пациента

*Программа разработки материала:*

1. Построение макетов статистических таблиц.
2. Целесообразно построить таблицу так, как приведено ниже

ФИО больного	Диагноз		Пол		Возраст		Цель посещения		Случай обслуживания	
	1	2	М	Ж	≤50	≥50	Лечение	Консультация	1-ый	2-ой
Иванов	+	-	+	-	+	-	-	+	-	+
Петрова	-	+	-	+	-	+	+	-	+	-

*План исследования:*

1. Объект – больной человек
2. Вид – выборочное
3. Время проведения – февраль 2006 года
4. Руководство и контроль – преподаватели курса социальной медицины

*Разработка материала – 25 талонов амбулаторного пациента:*

1. Проверка материала на наличие изучаемых признаков



2. Шифровка
3. Раскладка по шифрам
4. Подсчет
5. Занесение данных в таблицу

*Вывод:*

В соответствии с полученными данными среди больных преобладают женщины, большее количество из них в возрасте 30 – 55 лет. По цели посещения врача преобладает диагностика и лечение имеющейся патологии. Практически все случаи обращения повторные (более 85%). Среди отдельных форм заболеваний чаще других встречается гипертоническая болезнь.

Для выяснения причин, способствующих возникновению заболевания, необходимо дополнительно изучить условия жизни больных, условия труда, своевременность обращения за мед. помощью, а также вычислить производные величины, которые дадут возможность своевременно и объективно проанализировать материал, разработать необходимые меры профилактики.

### **Задачи для самостоятельного решения**

**1.** При изучении влияния аборта на младенческую смертность при последующих беременностях определены учетные признаки единиц наблюдения:

- Наличие аборта в анамнезе
- Количество аборт
- Состояние здоровья матери
- Гинекологические заболевания

Составить макет статистической таблицы «Причины младенческой смертности».

**2.** Составить план медико-статистического исследования согласно условию задачи №1.

**3.** Решить приведенные ниже тестовые ситуационные задачи. Мотивировать свой ответ.

В крупном промышленном регионе изучается эффективность деятельности санэпидстанции (СЭС). Необходимо определить единицу наблюдения.

- каждый случай инфекционного заболевания (учтенный СЭС)
- укомплектованность врачами гигиенистами и эпидемиологами
- каждая СЭС региона
- своевременность работы в очаге инфекционного заболевания

- качество проведения текущего санитарного надзора

В городе Н. было проведено социально гигиеническое исследование. В соответствии с программой исследования все население города сначала было разделено по месту жительства (порайонно), а потом из каждой группы пропорционально отбирались единицы наблюдения. Укажите метод формирования выборочной совокупности.

- механический отбор
- случайный
- парно-сопряженный
- когортный
- типологический

Виды статистического наблюдения по времени:

- Текущее
- Единовременное
- Гнездовое
- Выборочное
- Сплошного массива

Какие вопросы включает план статистического исследования?

- Формулировка цели исследования
- Группировка признаков
- Выбор объекта и единицы наблюдения
- Формулировка задач исследования
- Выбор места и времени исследования
- Выбор признаков, подлежащих регистрации
- Составление макетов таблиц

Требования к статистической информации:

- Качественная однородность
- Достоверность
- Сбор только путем переписей
- Сбор при помощи анкетирования
- Текущая регистрация событий
- Полнота сведений
- Своевременность поступления материалов

### **Контрольные вопросы**

1. Социальная гигиена как наука. История возникновения и развития.
2. Задачи социальной медицины и организации здравоохранения
3. Методы социальной медицины и организации здравоохранения, их использование для изучения здоровья населения и работы мед. учреждений
4. Медицинская статистика и ее разделы
5. Учетные документы и другие источники получения информации для решения задач медицинской статистики
6. Требования к статистической информации

7. Виды статистических совокупностей
8. Методы отбора единиц наблюдения для выборочной статистической совокупности

## **Относительные величины и их графическое изображение**

Относительные величины исчисляются при выполнении третьего этапа статистического исследования.

**Относительная величина** представляет собой результат сопоставления двух статистических показателей, дает цифровую меру их соотношения. Она получается путем деления сравниваемого показателя на другой показатель, принимаемый за базу сравнения.

Относительные величины делятся на две группы:

- относительные величины, полученные в результате соотношения одноименных статистических показателей;
- относительные величины, представляющие результат сопоставления разноименных статистических показателей.

К относительным величинам первой группы относятся: относительные величины динамики, относительные величины структуры (экстенсивные показатели), соотношения и наглядности.

Результат сопоставления одноименных показателей представляет собой краткое отношение (коэффициент), показывающее, во сколько раз сравниваемая величина больше (или меньше) базисной. Результат может быть выражен в процентах, показывая, сколько процентов сравниваемая величина составляет от базы.

**Относительные величины динамики** характеризуют изменение явления во времени. Они показывают, во сколько раз увеличился (или уменьшился) объем явления за определенный период времени, их называют коэффициентами роста. Коэффициенты роста можно исчислять в процентах, для этого отношения умножают на 100. Их называют темпами роста. Коэффициенты роста и темпы роста можно определять с переменной или постоянной базой.

Темпы роста с переменной базой получают при сравнении уровня явления каждого периода с уровнем предшествующего периода. Темпы роста с постоянной

базой сравнения получают путем сопоставления уровня явления в каждом отдельном периоде с уровнем одного периода, принятого за базу. Выбор базы сравнения нередко имеет существенное значение. Так, в ряде случаев в качестве базы сравнения принимаются годы, являющиеся исторически обусловленной границей отдельных периодов времени.

**y1 ; y2 ; y3 ; y4** — уровни явления за одинаковые последовательные периоды (например, уровень заболеваемости по годам дифтерией).

Темпы роста в процентах с переменной базой (цепные темпы роста):

$$T_{p_1} = \frac{y_2}{y_1} \cdot 100; \quad T_{p_2} = \frac{y_3}{y_2} \cdot 100; \quad T_{p_3} = \frac{y_4}{y_3} \cdot 100.$$

Темпы роста с постоянной базой (базисные темпы роста):

$$T'_{p_1} = \frac{y_1}{y_k} \cdot 100; \quad T'_{p_2} = \frac{y_2}{y_k} \cdot 100; \quad T'_{p_3} = \frac{y_3}{y_k} \cdot 100,$$

**Относительные величины структуры (экстенсивные показатели)** характеризуют долю отдельных частей в общем объеме совокупности и выражаются в долях единицы или в процентах. Каждую относительную величину структуры называют удельным весом. Примером такой относительной величины может быть уровень заболеваемости гломерулонефритом за текущий год по сравнению с другой нефрологической патологией.

**Относительные величины соотношения** отражают отношение численности двух частей единого целого, т. е. показывают, сколько единиц одной группы приходится в среднем на одну, на десять или на сто единиц другой группы изучаемой совокупности (например, сколько коек приходится на 10000 жителей района, сколько врачей приходится на 1000 человек, сколько детей приходится на 1 педиатра поликлиники).

**Относительные величины наглядности** отражают результаты сопоставления одноименных показателей, относящихся к одному и тому же периоду (или моменту)

времени, но к разным объектам или территориям (например, сравнивается годовая производительность труда по двум предприятиям или сравнивается загруженность койки в 2 стационарах).

Вторая группа относительных величин, представляющая собой результат сопоставления разноименных статистических показателей, носит название **относительных величин интенсивности (интенсивных показателей)**.

Они являются именованными числами и показывают итог числителя, приходящийся на одну, на десять, на сто единиц знаменателя.

В эту группу относительных величин включаются показатели производства продукции на душу населения; показатели потребления продуктов питания и непродовольственных товаров на душу населения; показатели, отражающие обеспеченность населения материальными и культурными благами; показатели, характеризующие первичную заболеваемость в районе:

$$\begin{array}{l} \text{Относительная} \\ \text{величина} \\ \text{структуры, \%} \end{array} = \frac{\text{Число единиц (или объем признака) по группе}}{\text{Общее число единиц (или объем признака) по всей совокупности}} \cdot 100.$$

**Графические изображения статистических данных** облегчают их обобщение и анализ. Графики применяются для характеристики развития явления во времени, в пространстве, отображения структуры явления и структурных сдвигов, при контроле за выполнением плана, изучении взаимосвязи между явлениями.

**По способу построения графики делятся на:**

- Диаграммы
- Картограммы
- Картодиаграммы

**Диаграмма** — изображение статистических данных при помощи геометрических фигур, линий, точек.

**Картограмма** — это географическая (контурная) карта, которая графически характеризует пространственное распределение какого-либо статистического показателя путем различной окраски, штриховки и т. д. (например, плотность населения в различных регионах).

**Картодиаграмма** - это совмещение картограммы с диаграммой, т. е. в отдельных районах условными знаками наносят абсолютные значения статистических показателей.

Самым распространенным видом графиков являются **диаграммы**, которые делятся на:

- Линейные
- Столбиковые
- Структурные
- Фигурные
- Знаки Варзара и др.

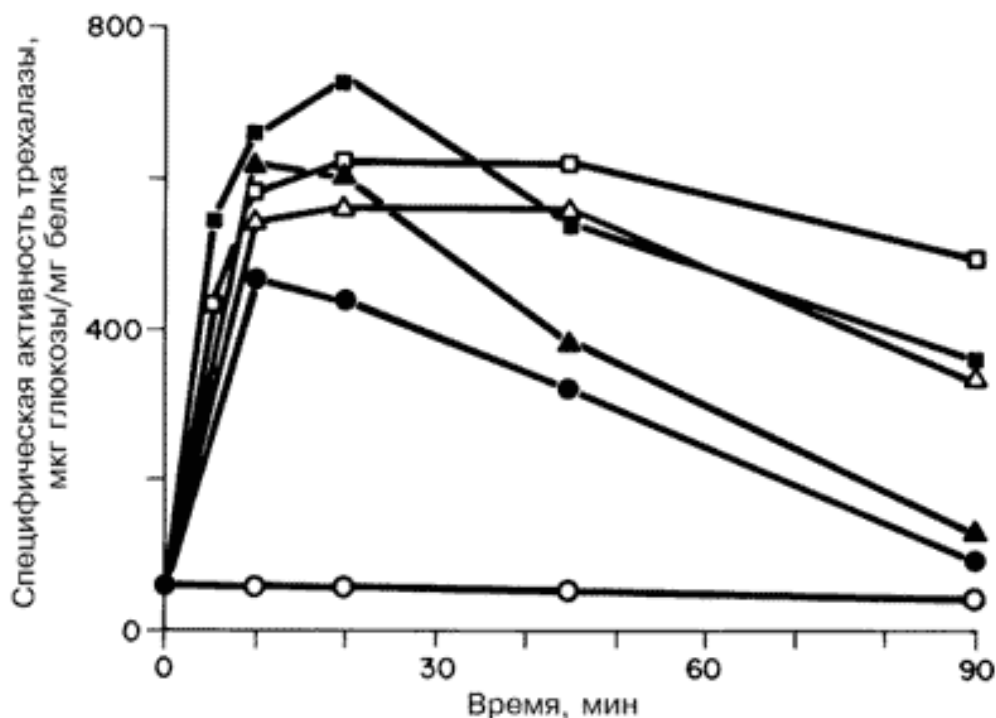
**Линейные диаграммы** — наиболее простой способ наглядного изображения статистических данных, когда изучаемое явление представляется в виде отрезков ломаной линии, называемой статистической кривой. Они применяются для характеристики и сравнения развития различных явлений во времени, пространстве, а также для отображения взаимосвязи между явлениями.

*Для построения линейной диаграммы* используется прямоугольная система координат. На оси абсцисс (по горизонтальной шкале) откладываются равные отрезки, представляющие собой периоды времени, на ось ординат наносят масштаб для отображения уровня явления. Соединение точек, построенных на координатной системе, дает ломаную линию, представляющую собой закономерность развития явления.

Рекомендуется строить координатную сетку с учетом соотношения масштабов по осям координат примерно 1 : 1,5 (правило «золотого сечения»), т. е. с учетом соотношения масштабов по сторонам занятого графиком пространства по вертикали и горизонтали.

*Преимуществом линейных графиков* является то, что на одном графике имеется возможность отображения закономерности нескольких явлений. Разновидностью линейных диаграмм являются контрольно-плановые

графики, обеспечивающие оперативный контроль за ходом выполнения задания как за отдельные промежутки (дни, пятидневки), так и нарастающим итогом с начала периода.

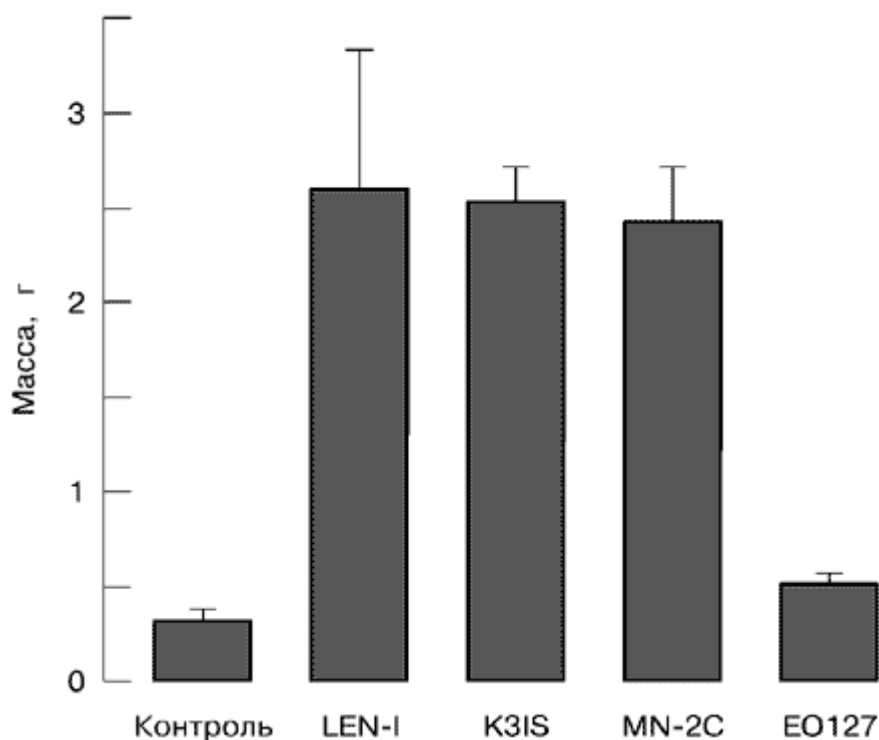


Линейный график. Шкалы по обеим осям (X и Y) линейные. По оси X отложена зависимая переменная (активность фермента трехазазы), а по оси Y – независимая переменная (время). Обратите внимание на то, что все метки делений находятся на равном расстоянии друг от друга и имеют одинаковую длину. Метки делений подписаны через общепринятые интервалы. Они хорошо видны, но не слишком велики. Нулевые отметки подписаны по обеим осям в точке их пересечения. Подписи меток оси Y расположены горизонтально, так же как и по оси X. Подпись оси Y идет параллельно оси снизу вверх, единицы измерения отделены от названия переменной. Черные квадратные маркеры точек максимально удалены от черных кружков. Маркеры в три раза больше, чем толщина линий. Вблизи черных маркеров линии разрываются. Контуры белых маркеров достаточно жирные. Кривые построены методом интерполяции и соединяют каждые две соседние точки на графике. При пересечении линии разрываются, за исключением одного пересечения, где угол, под которым линии пересекаются, не позволяет сделать разрыв. Различаются только маркеры, начертание линий одинаковое.

Для сравнения различных величин между собой и для изображения динамики могут быть использованы **столбиковые (ленточные) диаграммы**. Для их построения также используется система прямоугольных координат. Основания столбиков одинакового размера, представляющие собой периоды времени (годы, месяцы, дни), размещаются на оси абсцисс, а вершины столбиков



соответствуют величине изучаемого показателя. Столбиковые диаграммы называют ленточными, если столбики расположены горизонтально в виде лент.



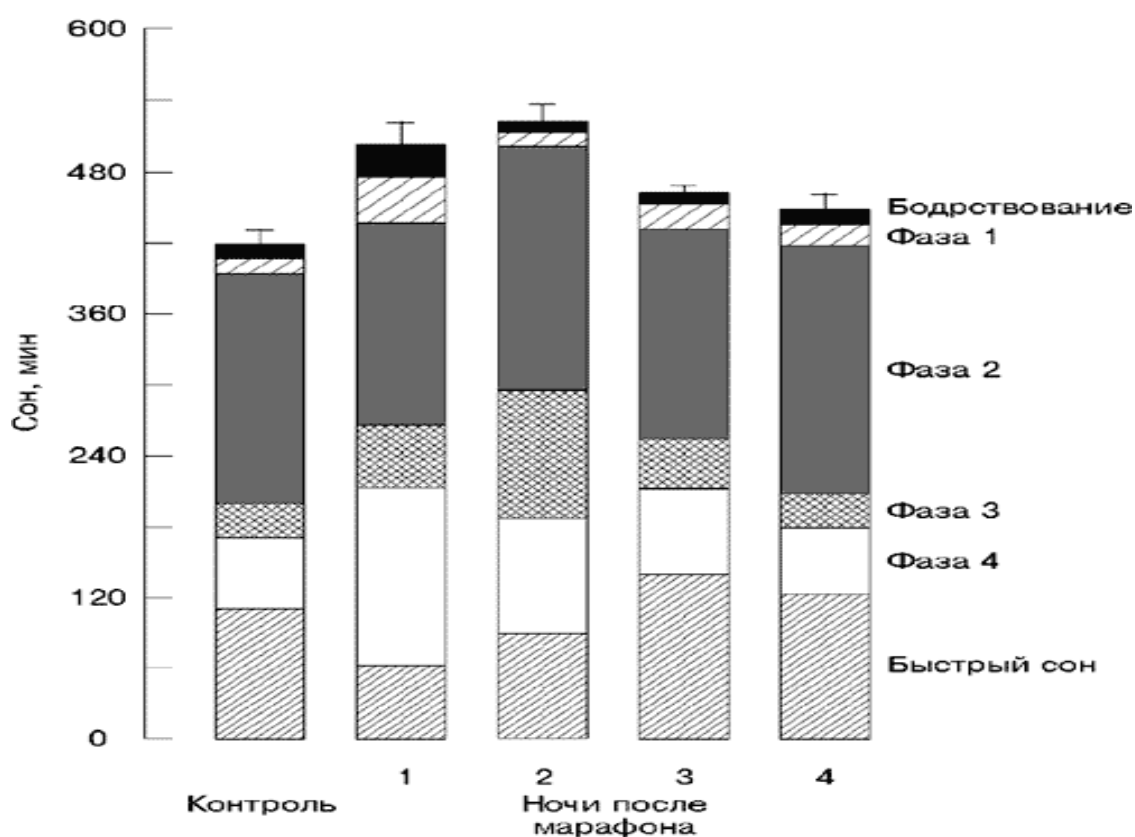
Столбиковая диаграмма с вертикальным расположением столбиков. Ось находится слева, начинается от нуля и ноль подписан. Основная линия не проведена, а намечена основаниями столбиков. Подписи к столбикам находятся ниже основной линии и каждая подпись выровнена по центру своего столбика. Все столбики одинаковой ширины и все промежутки между столбиками также одинаковой ширины; ширина столбиков больше, чем ширина промежутков. Контуры столбиков жирные, заливка – темно-серая. Столбик, представляющий контрольную группу, находится слева. Линии ошибок центрированы по отношению к своим столбикам и направлены вверх. Представлена верхняя половина линий ошибок, так как заливка столбиков темная.

**Секторные (структурные) диаграммы** применяются для изображения структуры явления и характеристики структурных сдвигов. При построении таких графиков состав совокупности выражается относительными величинами структуры, исчисленными в процентах. Они могут быть двух видов: столбиковые и круговые. Общая высота столбика и площадь круга отображают целое и принимаются соответственно за 100%.

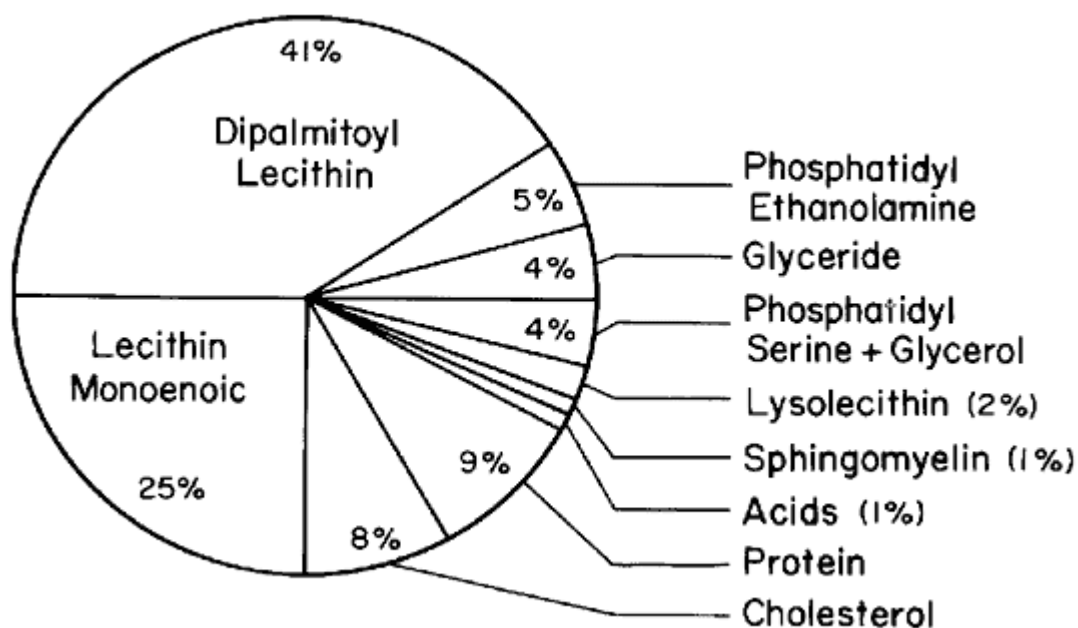
При построении круговой диаграммы необходимо проценты перевести в градусы, учитывая, что каждый

процент равен  $3,6^\circ$  ( $360 : 100$ ). Для тех же целей может быть использована и **внутристолбиковая диаграмма**.

Эти виды графиков показывают соотношения частей целого. На внутристолбиковой диаграмме части закрашивают по-разному (заливками и штриховками). На секторной диаграмме части показывают секторами разной величины. Если сопоставляют несколько целых, то используют внутристолбиковую диаграмму.

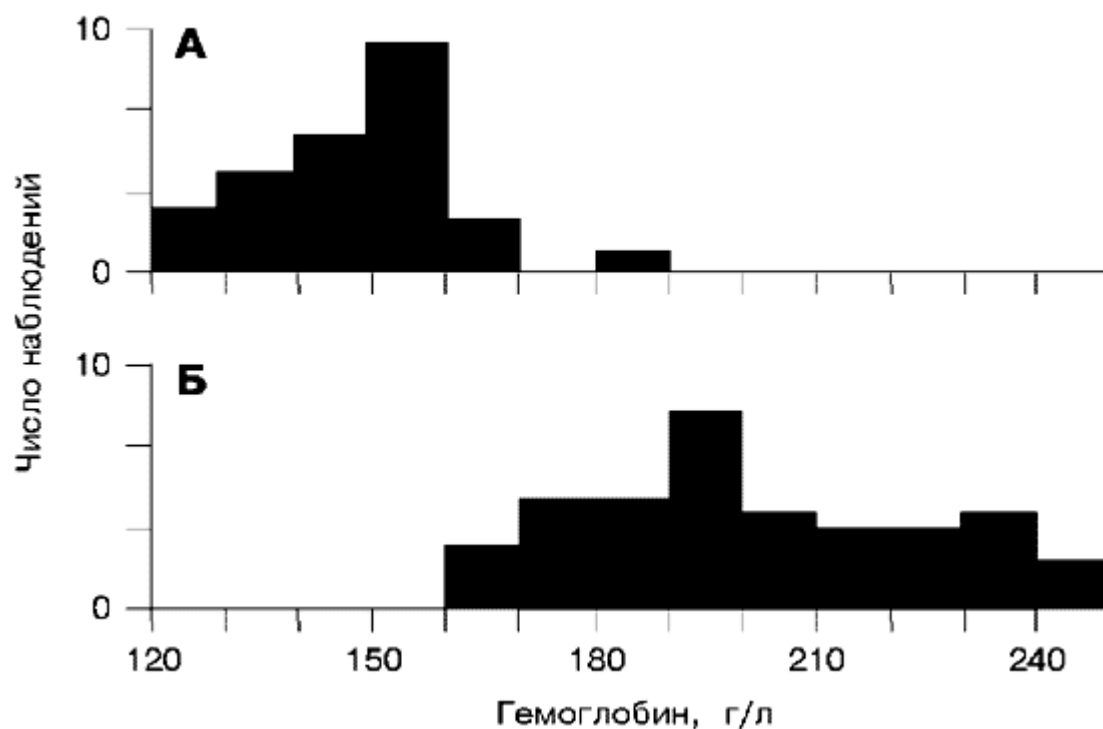


Составная столбиковая диаграмма. Компоненты различаются штриховкой и заливкой, которые обозначены подписями в правой части рисунка. Составная столбиковая диаграмма в данном случае удобнее, чем секторная, так как позволяет сравнивать размеры компонентов в соседних столбиках.

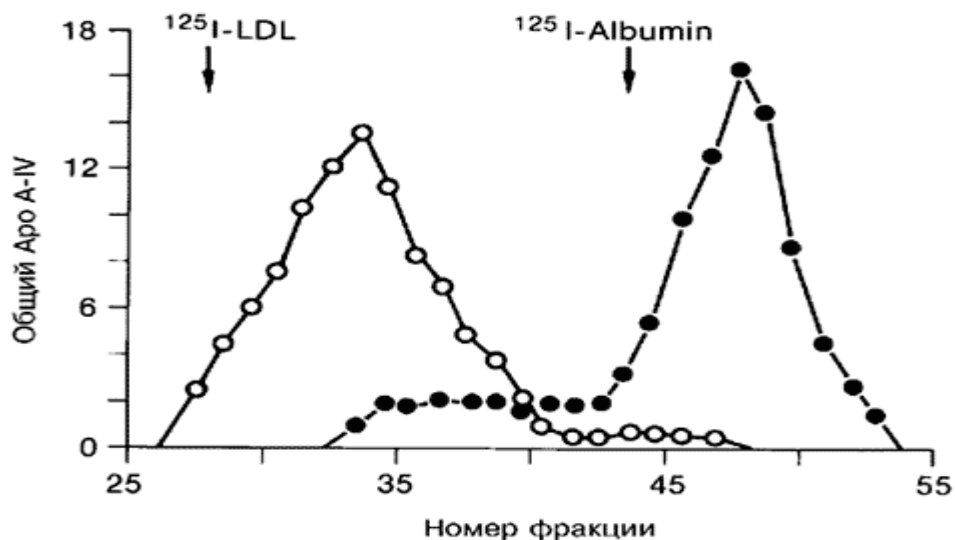


Секторная диаграмма. Проценты указаны в каждом секторе, сумма процентов составляет 100. Проценты вписаны рядом с краем круга. Подписи отделены от процентов и центрированы по отношению к процентам. В мелких секторах указаны только проценты – рядом с наружным краем, подписи же вынесены вправо и расположены столбцом. Направляющие линии показывают, к какому сектору относится каждая подпись. Если сектор совсем узкий, то и проценты и подпись выносят за пределы круга. Все проценты и подписи расположены горизонтально.

**Гистограммы и полигоны частот** показывают частотные распределения на координатной сетке. Частотное распределение показывает общее число наблюдений в каждом интервале распределения непрерывной величины, например времени или массы. На гистограмме частотное распределение показывают путем изображения последовательности прямоугольников. Ширина каждого прямоугольника соответствует диапазону значений, принимаемых одной переменной на оси X. Высота прямоугольника соответствует наблюдаемой частоте, которая откладывается на оси Y. Площадь гистограммы показывает распределение данных. Гистограмма изображает распределение только в том случае, если ширина всех прямоугольников одинаковая. Гистограмму можно изображать с вертикальными линиями, обозначающими границы прямоугольников, или без них. Гистограмму можно заливать черным или темно-серым цветом или оставлять белой.



Две гистограммы. Каждый прямоугольник соответствует одному интервалу распределения непрерывной величины. Высота каждого прямоугольника показывает частоту в соответствующем интервале. Поскольку ширина всех прямоугольников равна, то и высоты прямоугольников, и их площади отражают частотное распределение. Не следует накладывать одну гистограмму на другую, а нужно поместить их на разных осях. Обратите внимание на то, что ось  $Y$  начинается с нуля. Буквами **A** и **B** отмечены две части сложного рисунка, эти буквы – самые жирные на графике. На сложном рисунке подписывается только одна ось  $Y$  и только одна ось  $X$ . Подписи меток делений требуются также только на одной оси  $X$ .



Два полигона частот наложены друг на друга для сравнения перекрывающихся распределений. Каждая вершина многоугольника (каждый маркер) соответствует частоте в определенном интервале. Площадь каждого многоугольника соответствует частотному распределению. Ось Y начинается с нуля и стороны обоих многоугольников продлены от первых и последних вершин до пересечения с нулевой линией. Стрелки изображены для того, чтобы отметить на многоугольнике точки, представляющие особый интерес.

**На полигонах частот** распределение изображено не прямоугольниками, а средними точками на каждом интервале оси X, соединенными между собой линиями. Эти линии продолжаютсЯ до пересечения с основной линией (частота 0). Такой способ удобен для изображения перекрывающихся распределений, что неудобно делать с помощью гистограмм.

**Знаки Варзара** (по имени статистика В. Е. Варзара) являются разновидностью столбиковых диаграмм. Они позволяют отобразить на графике сложное явление, представляющее собой произведение двух показателей. Например, эффективность лечения — произведение производительности труда мед. персонала и применяемых методов лечения. Если в прямоугольнике одну сторону взять пропорционально уровню производительности труда, а другую — пропорционально численности работников, то площадь прямоугольника будет пропорциональна объему продукции.

**Круговые (радиальные) диаграммы** – используются чаще для отображения внутригодовой динамики какого-либо показателя. Например, рождаемость по месяцам в городе или районе. На радиусах, проведенных через 30°, откладываются отрезки определенной длины, пропорциональной величине рождаемости за данный месяц.

### **Задача – эталон №1**

**Условие.** В районе А зарегистрировано 500 случаев инфекционных заболеваний. Из них: эпидемический паротит – 60, корь – 100, других – 340. Определите относительные величины. Изобразите графически полученные данные.

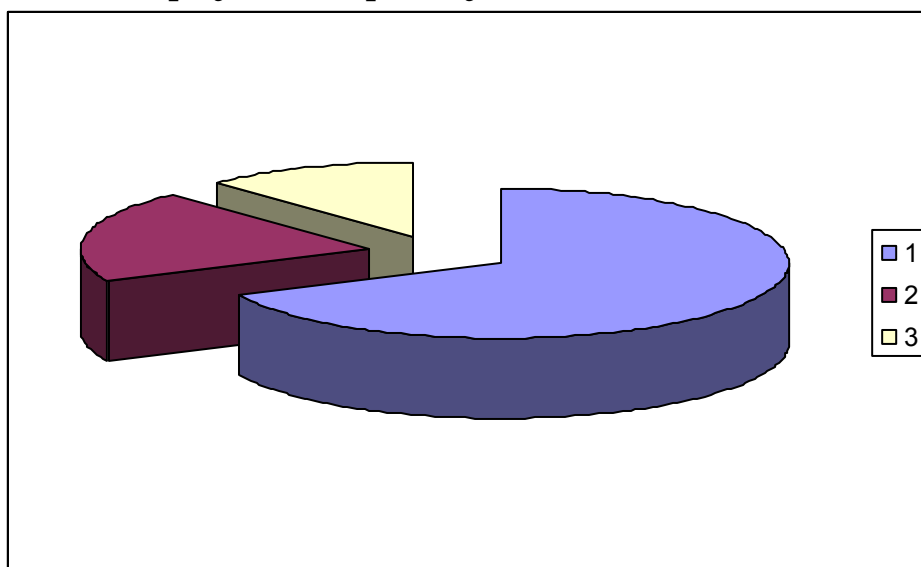
**Решение.** Поскольку имеем статистическую совокупность и ее части, ею не продуцируемые, то говорить нужно о экстенсивном показателе. Определим его для каждого из приведенных в условии заболеваний.

$$\text{ЭП}_{\text{паротита}} = \frac{60}{500} \times 100\% = 12\%$$

$$\text{ЭП}_{\text{кори}} = \frac{100}{500} \times 100\% = 20\%$$

$$\text{ЭП}_{\text{другие}} = \frac{340}{500} \times 100\% = 68\%$$

Графически экстенсивные показатели, как показатели структуры явления, наиболее целесообразно отображать с помощью секторной диаграммы. Построим по имеющимся данным секторную диаграмму.



1 – другие инфекционные заболевания, 2 – корь, 3 – паротит.

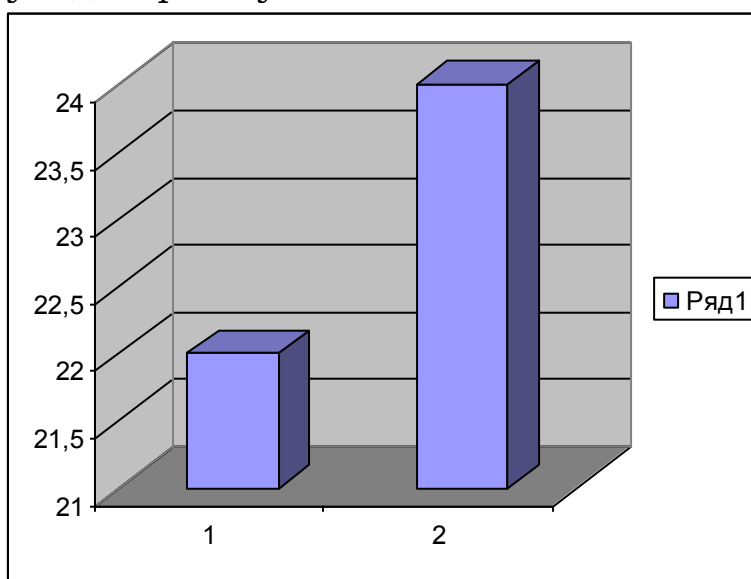
## Задача – эталон №2

**Условие.** Рассчитать показатель наглядности, если известно, что госпитализировано в 2005 году 24% обратившихся, а в 2006 – 22%. Исходные данные отобразить графически.

**Решение.** Рассчитываем показатель наглядности по формуле, исходя из условия.

$$\text{ПН} = \frac{22}{24} \times 100\% = 91\%$$

Графически ежегодный прирост (уменьшение) какого-либо явления наиболее целесообразно отображать в виде линейной или столбиковой диаграммы. Построим столбиковую диаграмму.



1 – уровень госпитализации за 2005 год, 2 – за 2006 год.

## Задачи для самостоятельного решения

**1.** В городе 120 000 человек. Определите показатель рождаемости, если в предыдущем году родилось 1200 детей. Классифицируйте полученный показатель. Отобразите графически динамику за последние 5 лет, если рождаемость по годам была следующей: 2006 г. – ?, 2005 г. – 0,02%, 2004 г. – 0,013%, 2003 г. – 0,09%, 2002 г. – 0,011%.

**2.** На 120 000 населения в городских больницах 300 терапевтических коек. Определите обеспеченность населения койками. Классифицируйте полученный показатель.

**3.** Решите приведенные ниже ситуационные задачи. Обоснуйте Ваши ответы.

С целью анализа состояния здоровья населения и уровня медпомощи в кардиологическом диспансере были рассчитаны следующие показатели: уровень первичной заболеваемости болезнями системы кровообращения - 62%; уровень общей заболеваемости болезнями системы кровообращения - 483,55 ‰; уровень смертности от болезней системы кровообращения - 10,9 ‰; удельный вес случаев смерти от болезней системы кровообращения среди всех причин смерти - 67,0%; уровень первичной инвалидности от болезней системы кровообращения - 16,2 на 10 тыс. населения. Какой из приведенных показателей является экстенсивной величиной?

- а. уровень первичной заболеваемости болезнями системы кровообращения
- б. удельный вес случаев смерти от болезней системы кровообращения среди всех причин смерти
- в. уровень общей заболеваемости болезнями системы кровообращения
- г. уровень смертности от болезней системы кровообращения
- д. уровень первичной инвалидности от болезней системы кровообращения

Для расчета показателей наглядности нужны следующие исходные данные:

- а. Часть явления, все явление
- б. Все явление, среда
- в. Среда, все явление, основание (100, 1000, 10000)
- г. Часть явления в ‰, все явление
- д. Явление, принятое за единицу, сравниваемое явление.

Выберите метод графического изображения структуры причин материнской смертности:

- а. линейная диаграмма.
- б. столбиковая диаграмма.
- в. объёмная диаграмма.
- г. секторная диаграмма.
- д. картодиаграмма.

К какому виду относительных величин относится показатель обеспеченности населения врачами?

- а. интенсивный.
- б. соотношения.
- в. экстенсивный.
- г. наглядности.
- д. гармоничности.

Виды относительных величин:

- а. Инвалидности, интенсивности, экстенсивности
- б. Соотношения, интенсивности, экстенсивности
- в. Заболеваемости, интенсивности, экстенсивности
- г. Смертности, интенсивности, экстенсивности
- д. Летальности, интенсивности, экстенсивности

### **Контрольные вопросы**

1. Относительные величины и их виды. Методы их определения
2. Вычисление интенсивного показателя



3. Вычисление экстенсивного показателя
4. Вычисление показателя соотношения
5. Вычисление показателя наглядности
6. Примеры использования относительных показателей в практике здравоохранения
7. Для чего используются графические изображения?
8. Виды графиков
9. Основные правила построения графиков
10. Выбор вида графика
11. Методы построения графиков

## **Вариационные ряды. Средние величины. Методика их вычисления и оценки.**

Кроме относительных величин (коэффициентов), характеризующих частоту (интенсивность), либо состав изучаемого явления, статистические совокупности с их количественной стороны могут быть охарактеризованы при помощи средних величин. При изучении физического развития населения, закономерностей течения различных процессов в здоровом и больном организме, для оценки эффективности лекарственных препаратов и решения целого ряда других задач используются средние величины.

**Средние величины** получаются из рядов распределения (вариационных рядов). В таком ряду количественно изменяющийся признак носит название варьирующего, а отдельные его количественные выражения называются **вариантами**.

Числа, показывающие, как часто встречается та или иная варианта в составе данного ряда, носят название **частот**.

Ряд, в котором сопоставлены варианты и соответствующие этим вариантам частоты и который, следовательно, показывает распределение изучаемой совокупности по величине какого-либо варьирующего признака, носит название **вариационного ряда** (распределение призывников по росту, новорожденных по весу и т.п.).

### **Выделяют формы вариационного ряда:**

- простой (частота встречаемости каждой из вариантов равна 1)
- взвешенный (частоты встречаемости каждой из вариантов написаны под значением варианты)
- ранжированный ряд
- дискретный ряд
- интервальный ряд (применяется при очень малой разнице между значениями вариант и высокой  $n$  – вариационный ряд разбивается на несколько интервалов – меньше 100, 100-200, 200-300, более 300 (открытый) или 0-100, 100-200, 200-300, 300-400 (закрытый))

**Среднее арифметическое, или просто среднее,** — одна из основных характеристик выборки. Среднее арифметическое – такое значение признака, сумма отклонений от которого выборочных значений признака равна нулю (с учетом знака отклонения).

Среднее принято обозначать той же буквой, что и варианты выборки, с той лишь разницей, что над буквой ставится символ усреднения — черта. Например, если обозначить исследуемый признак через  $X$ , а его числовые значения — через  $x_i$ , то среднее арифметическое имеет обозначение  $\bar{x}$ .

Среднее арифметическое, как и другие числовые характеристики выборки, может вычисляться как по необработанным первичным данным, так и по результатам группировки этих данных.

Для несгруппированных данных среднее арифметическое определяется по следующей формуле:

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} = \frac{1}{n} \sum_{i=1}^n x_i$$

где  $n$  — объем выборки;  $x_i$  — варианты выборки.

Если данные сгруппированы, то формула приобретает вид:

$$\bar{x} = \frac{n_1 x_1 + n_2 x_2 + \dots + n_k x_k}{n} = \frac{1}{n} \sum_{i=1}^k n_i x_i$$

где  $n$  — объем выборки;  $k$  — число интервалов группировки;  $n_i$  — частота  $i$ -ого интервала;  $x_i$  — срединное значение  $i$ -ого интервала.

Нахождение среднего арифметического непрерывного вариационного ряда осложняется если крайние интервалы не замкнуты (то есть имеют вид “менее 10” или “более 60”). В этом случае считается, что ширина первого интервала равна ширине второго, а ширина последнего — ширине предпоследнего.

Среднее арифметическое, вычисленное по второй формуле, называют также взвешенным средним, подчеркивая этим, что в формуле  $x_i$ , суммируются с коэффициентами (весами), равными частотам попадания в интервалы группировки.

**Медиана и мода.** Вариационный ряд, кроме средней арифметической, может быть охарактеризован еще двумя величинами — медианой и модой.

**Модой** (обозначается  $M_o$ ) называется та варианта, которой соответствует наибольшее количество частот вариационного ряда. Если мода находится в интервале, например 149,0 – 152,9, с некоторым приближением можно принять моду равной середине этого интервала, т.е.  $M_o = 151,0$  см.

В мед. статистике применение моды довольно ограничено. Модой можно пользоваться для оценки средней длительности заболеваний, особенно при малом количестве больных данной болезнью.

В том случае несколько больных с особо длинными или очень короткими сроками лечения окажут значительное влияние на величину средней арифметической величины, если воспользоваться ею для определения средней длительности заболевания. Здесь мода, т.е. обычная длительность заболевания, окажется более полезной для практического использования.

**Медианой** (обозначается  $M_e$ ) называется варианта, делящая вариационный ряд на две равные половины, т.е. та варианта, занимающая срединное положение в вариационном ряду.

Медиана применяется в санитарной статистике относительно редко. С помощью медианы определяется, например, так называемая вероятная продолжительность предстоящей жизни в таблицах смертности.

*Основным отличием медианы и моды от средней арифметической является то, что на их размеры не оказывают влияния величина крайних значений вариантов, имеющих в вариационном ряду, тогда как при определении средней арифметической принимаются во внимание значения всех вариантов.*

**Среднее квадратическое отклонение.** Еще одним параметром вариационного ряда (величиной, характеризующей вариационный ряд) является среднее квадратическое отклонение, обозначаемое  $\sigma$  (малая греческая сигма).

Среднее квадратическое отклонение равняется квадратному корню из суммы произведений частот вариационного ряда на квадраты отклонений вариантов от средней арифметической, деленной на число частот:

$$\sigma = \sqrt{\frac{\sum p d^2}{n}}$$

Более правильно в знаменателе подкоренного выражения ставить не  $n$ , а  $n-1$ . При достаточно большом количестве наблюдений уменьшение знаменателя на 1 не сказывается сколько-нибудь существенно на результате, так как после всех вычислений изменяется только величина второго или даже третьего десятичного знака. Однако при малом количестве наблюдения (примерно 30 и

менее), что почти всегда имеет место при статистической обработке клинических и лабораторно-экспериментальных материалов, это уточнение имеет значение. В этих случаях:

$$\sigma = \sqrt{\frac{\sum p d^2}{n - 1}}$$

*Вычисление среднего квадратического отклонения по амплитуде.* Если отсутствуют необходимые исходные данные для вычисления среднего квадратического отклонения обычным путем, может быть использован приближенный способ вычисления  $\sigma$  по амплитуде вариационного ряда.

**Амплитудой ряда** называется разность между наибольшей и наименьшей вариантами ( $v_{\max} - v_{\min}$ ).

Среднее квадратическое отклонение, исчисляемое по амплитуде, несколько отличается по величине от  $\sigma$ , вычисленной обычными способами. Различие это тем больше, чем больше число наблюдений, использованных для составления вариационного ряда. Поэтому определение среднего квадратического отклонения по амплитуде более целесообразно производить преимущественно при ориентировочных расчетах. Вычисление производится по формуле:

$$\sigma = \frac{ampl}{k} = \frac{v_{\max} - v_{\min}}{k},$$

**Значение среднего квадратического отклонения.** Средняя арифметическая характеризует одной величиной весь вариационный ряд. Однако, чем больше варьирует индивидуальные значения вариантов, тем, очевидно, менее точно характеризуется вариационный ряд средней арифметической.

Ряд с большей амплитудой имеет большее среднее квадратическое отклонение (амплитуда приблизительно равно  $6\sigma$ ).

Следовательно, две одинаковые средние, полученные из вариационных рядов с различной амплитудой, не в одинаковой степени характеризуют свои ряды. Та из них, которая имеет меньшее среднее квадратическое отклонение и, следовательно, получена из вариационного ряда с меньшей вариабельностью, своим размером

будет больше приближаться к действительной величине значительного большинства единиц ряда.

Отклонение размера роста данного лица от средней величины роста всего коллектива, однородного в отношении пола, возраста, этнической и социальной принадлежности и пр., не превышающее среднего квадратического отклонения, считается находящимся в пределах нормы. Отклонение (в любую сторону) больше чем на  $1\sigma$ , но меньше чем на  $2\sigma$  считается субнормальным, а отклонение больше, чем на  $2\sigma$  – значительно выше или ниже среднего.

*Среднее квадратичное отклонение может быть применено для:*

- а) для суждения о размахе вариационных рядов и сравнительной оценки типичности (представительности) средних арифметических величин. Это необходимо в дифференциальной диагностике при определении устойчивости признаков;
- б) для реконструкции вариационного ряда, т.е. восстановления его частотной характеристики на основе *правила «трех сигм»*. В интервале  $M+3\sigma$  находится 99,7% всех вариант ряда, в интервале  $M+2\sigma$  — 95,5% и в интервале  $M+1\sigma$  — 68,3% вариант ряда;
- в) для выявления «выскакивающих» вариант (при сопоставлении реального и реконструированного вариационных рядов);
- г) для определения параметров нормы и патологии с помощью сигмальных оценок;
- д) для расчета коэффициента вариации;
- е) для расчета средней ошибки средней арифметической величины.

**Коэффициент вариации.** Стандартное отклонение выражается в тех же единицах измерения, что и характеризуемый им признак. Если требуется сравнить между собой степень варьирования признаков, выраженных в разных единицах измерения, возникают определенные неудобства. Пусть, например, результаты в беге на 100 м, показанные группой IX классов, имеют стандартное отклонение 0,9 с (при среднем времени 14,8 с), а исследование роста тех же учащихся показывает, что его стандартное отклонение составляет 6 см (при среднем росте 168 см). Какой из признаков варьирует сильнее? Очевидно, что только на основании сравнения стандартных отклонений на этот вопрос ответить нельзя. Требуется сопоставить стандартные отклонения со средними арифметическими этих признаков. Поэтому вводится относительный показатель называемый коэффициентом вариации.

$$V = \frac{\sigma_B}{\bar{x}_B}$$

Обычно он выражается в процентном отношении:

$$V = \frac{\sigma_B}{\bar{x}_B} \cdot 100 \%$$

Коэффициент вариации является относительной мерой рассеяния признака.

Коэффициент вариации используется и как показатель однородности выборочных наблюдений. Считается, что если коэффициент вариации не превышает 10%, то выборку можно считать однородной, т. е. полученной из одной генеральной совокупности.

Считается, что если:

1. коэффициент вариации менее 10%, то разнообразие признаков в ряду низкое
2. от 10% до 20% - то разнообразие признаков среднее
3. более 20% - разнообразие признаков высокое.

То есть можно говорить о том, что если коэффициент вариации менее 10% и разнообразие признаков низкое, то в большинстве случаев может быть применена средняя арифметическая вместо среднего квадратичного отклонения.

Практически коэффициент вариации применяется в основном для сравнения выборок из однотипных генеральных совокупностей.

К использованию коэффициента вариации нужно подходить с осторожностью. Продемонстрируем возможные ошибки на следующем примере. Если на основании многолетних наблюдений среднее арифметическое среднесуточных температур 8 марта составляет в какой-либо местности 0°C, то по формуле получим бесконечный коэффициент вариации независимо от разброса температур. Поэтому в данном случае коэффициент вариации не применим!

*Коэффициент вариации применим в следующих случаях:*

- а) для оценки разнообразия каждого конкретного вариационного ряда и, соответственно, суждения о типичности отдельной средней (т.е. ее способности быть полноценной обобщающей характеристикой данного ряда). Сильное разнообразие ряда свидетельствует о малой представительности (типичности) соответствующей средней величины и, следовательно, о нецелесообразности ее использования в практических целях.

б) для сравнительной оценки разнообразия разноименных вариационных рядов и выявления более-менее стабильных признаков, что имеет значение в дифференциальной диагностике.

### **Задачи для самостоятельного решения**

**1.** Решите приведенные ниже ситуационные задачи. Обоснуйте Ваш ответ.

Каким свойством статистической совокупности обусловлена необходимость вычисления средних величин?

- А. распределением признака.
- В. разнообразием признака.
- С. репрезентативностью признака.
- Д. средним уровнем признака.
- Е. взаимосвязью признаков.

Какое свойство средней арифметической используется для проверки правильности ее вычисления?

- А. средняя занимает срединное положение.
- В. средняя является обобщающей величиной.
- С. средняя имеет ту же размерность, что и варианты.
- Д. сумма отклонений вариантов от средней равна нулю.
- Е. сумма отклонений вариантов от средней равна единице.

Назовите формулу для расчета средней арифметической взвешенной:

- А.  $C = G/M \cdot 100$ .
- В.  $M = \sum VP/n$ .
- С.  $M = \sum V/n$ .
- Д.  $M = A \cdot \sum vp/n$ .
- Е.  $M = \sum V/n \cdot \sum ap/n$ .

Каково назначение средних величин?

- А. характеризуют частоту явления.
- В. характеризуют распределение целого на части.
- С. характеризуют числовые соотношения независимых величин.
- Д. характеризуют общую меру явления.
- Е. характеризуют динамику процесса.

Назовите элементы вариационного ряда:

- А. мода, медиана.
- В. варианта, частота.
- С. средняя арифметическая.
- Д. квадрат отклонения.
- Е. единица наблюдения, мода, отклонение.

От чего зависит величина ошибки (m)?

- А. числа наблюдений.
- В. способа выборки.
- С. изменчивости признака.
- Д. способа вычисления ошибки.
- Е. порядкового номера варианты.

Какой критерий применяется для оценки степени разнообразия признака в совокупности?

- А. коэффициент вариации.
- В. амплитуда.
- С. среднее квадратическое отклонение.
- Д. лимит.



Е. коэффициент вариации.

Какой из критериев используется для определения типичности средней арифметической?

- А. лимит.
- В. амплитуда.
- С. коэффициент вариации.
- Д. коэффициент корреляции.
- Е. среднее квадратическое отклонение.

Средний рост новорожденных мальчиков равен 50, 9 см при сигме 1, 66, а средняя масса тела – 3432 г. при сигме 500. По каким критериям можно сравнить степень вариабельности этих признаков?

- А. сигма.
- В. лимит.
- С. амплитуда.
- Д. коэффициент вариации.
- Е. коэффициент ассоциации.

Какой критерий применяется для оценки разнообразия признака в совокупности?

- А. лимит.
- В. амплитуда.
- С. среднее квадратическое отклонения.
- Д. Хи - квадрат.
- Е. коэффициент вариации.

Назовите формулу для оценки достоверности разности средних величин:

- А.  $M_{ген} = M_{выб} \pm m_t$ .
- В.  $P_{ген} = P_{выб} \pm m_t$ .
- С.  $M = \Sigma VP / n$
- Д.  $t = (M_1 - M_2) / \sqrt{m_1^2 + m_2^2}$ .
- Е.  $t = (P_1 - P_2) / \sqrt{m_1^2 + m_2^2}$ .

Формула для определения ошибки средней арифметической величины?

- А.  $M = \Sigma VP / n$ .
- В.  $M_{ген} = M_{выб} \pm m_t$ .
- С.  $m = \pm \sigma / \sqrt{n}$ .
- Д.  $\sigma = \pm \sqrt{p(d \cdot d) / (n - 1)}$ .
- Е.  $E = (M_1 - M_2) / \sqrt{m_1^2 + m_2^2}$ .

Какая минимальная величина разности сравниваемых величин принята в медицине при расчете критерия Стьюдента?

- А. 2.
- В. 1, 6.
- С. 1, 0.
- Д. 0, 6.
- Е. 1, 8.

При анализе использования коечного фонда (среднегодовая занятость койки, средняя продолжительность пребывания больного на больничной койке) в многопрофильной больнице города К. установлено, что учетные признаки имеют асимметрическое распределение. Какие статистические данные необходимо использовать для определения достоверности разницы этих показателей в разных отделениях?

- А. Параметрические
- В. Непараметрические
- С. Традиционные
- Д. Системного анализа
- Е. Все перечисленное

В населенном пункте К. изучается распространенность злокачественных новообразований. Получен результат  $P \pm 2m$ . Необходимо указать степень вероятности прогноза для полученного результата.

- А. 27,8%
- В. 68,0%
- С. 95,5%
- Д. 99,7%
- Е. 100,0%

Средняя продолжительность лечения больных гипертонической болезнью в больнице №1 составляла 17 суток ( $m=\pm 1,0$ ), а в больнице №2 - 15 суток ( $m=\pm 0,5$ ). Достоверна ли разность показателей?

- |   |  |
|---|--|
| A. Да                                     | D. Необходимо уменьшить число наблюдений.      |
| B. Нет                                    | E. Необходимо определить коэффициент вариации. |
| C. Необходимо увеличить число наблюдений. |  |

Средний рост мальчиков 6 лет составляет 116,9 см ( $\sigma = (\pm 4,2)$  см), средняя масса тела - 22,2 кг ( $\sigma = (\pm 3,4)$  кг). С помощью какого показателя можно сравнить степень варибельности признаков?

- |                             |  |
|-----------------------------|--|
| A. По коэффициенту вариации | D. По коэффициенту корреляции          |
| B. По дисперсии             | E. По средним квадратичным отклонениям |
| C. По амплитуде рядов       |  |

Из 1000 детей, которые были привиты от гриппа, заболело 100, а из 400 непривитых детей заболело гриппом 50. Какая из ниже перечисленных методик является наиболее приемлемой для проверки эффективности прививок от гриппа?

- |   |  |
|---|--|
| A. Вычисление коэффициента корреляции.    | C. Проверка точности вычисления показателей.             |
| B. Оценка меры варибельности показателей. | D. Оценка достоверности различий заболеваемости гриппом. |
|   | E. Метод стандартизации                                  |

Информация, которая не нужна для вычисления среднего квадратичного отклонения при расчете средней арифметической взвешенной:

- |                         |   |
|-------------------------|---|
| A. частота вариант      | D. все перечисленное верно              |
| B. число наблюдений     | E. отклонение любой варианты от средней |
| C. коэффициент вариации |   |

В медицинских и биологических исследованиях разница считается статистически достоверной при значении t:

- |            |              |                        |
|------------|--------------|------------------------|
| A. более 3 | C. более 1,5 | E. более или равно 0,5 |
| B. более 2 | D. более 1   |                        |

Частота осложнений при кишечной непроходимости среди прооперированных детей составила  $(3,0\pm 0,5)\%$ . Необходимо указать в каких пределах может измениться этот показатель при вероятности статистического прогноза в 95,5 %.

- |              |              |              |
|--------------|--------------|--------------|
| A. 3,0-0,5 % | C. 3,5-2,5 % | E. 4,5-1,5 % |
| B. 3,0-1,0 % | D. 4,0-2,0 % |              |

Средний рост новорожденных мальчиков равняется 50,9 см при сигме 1,66, а средняя масса тела - 3432 г при сигме 5,00. По какому критерию можно сравнить степень варибельности этих признаков?

- A. Сигма.
- B. Лимит.
- C. Амплитуда.
- D.  $t$  критерий Стьюдента
- E. Коэффициент вариации

### **Контрольные вопросы**

1. Вариационный ряд. Элементы вариационного ряда
2. Виды вариационных рядов. Правила построения рядов
3. Область применения средних величин в здравоохранении
4. Виды средних величин
5. Формулы расчета средней арифметической. Применение. Виды. Значение
6. Среднее квадратичное отклонение. Формулы. Применение. Значение
7. Коэффициент вариации. Формулы. Значение. Область применения. Оценка полученных данных
8. Метод сигмальных отклонений и его применение в практическом здравоохранении

## **Оценка достоверности полученных результатов. Методы. Задачи. Область применения**

В практической и научно-практической работе врачи обобщают результаты, полученные, как правило, на выборочных совокупностях. Для более широкого распространения и применения полученных при изучении репрезентативной выборочной совокупности данных и выводов надо уметь по части явления судить о явлении и его закономерностях в целом.

Учитывая, как правило, что врачи проводят исследования на выборочных совокупностях, теория статистики позволяет с помощью математического аппарата (формул) переносить данные с выборочного исследования на генеральную совокупность. При этом врач должен уметь не только пользоваться математическими формулами, но и делать выводы, соответствующие каждому способу оценки достоверности полученных данных. С этой целью врач должен знать способы оценки достоверности.

Применяя метод оценки достоверности результатов исследования для изучения общественного здоровья и деятельности учреждений здравоохранения, а также в своей научной деятельности, исследователь должен уметь правильно выбрать способ данной оценки. Среди методов оценки достоверности различают параметрические и непараметрические методы.

**Параметрическими** называют количественные методы статистической обработки данных, применение которых требует обязательного знания закона распределения изучаемых признаков в совокупности и вычисления их основных параметров.

В тех случаях, когда имеется малое количество наблюдений и характер распределения неизвестен, когда кроме количественных характеристик, результаты выражаются полуколичественными, а иногда описательными характеристиками (тяжесть заболевания, интенсивность реакции, результаты лечения), параметрические методы становятся непригодными. В этих ситуациях следует использовать непараметрические методы оценки достоверности.

**Непараметрическими** являются количественные методы статистической обработки данных, применение которых не требует знания закона распределения изучаемых признаков в совокупности и вычисления их основных параметров.

В то же время следует отметить, что назначение применения непараметрических методов гораздо шире, чем только оценка достоверности результатов исследования (в том числе они применяются и для характеристики одной выборочной совокупности, и для изучения связи между явлениями). В данном случае акцент сделан на оценке достоверности результатов исследования, как одном из наиболее важных разделов статистического анализа, поэтому непараметрические методы не представлены отдельной темой.

Как параметрические, так и непараметрические методы, используемые для сравнения результатов исследований, т.е. для сравнения выборочных совокупностей, заключаются в применении определенных формул и расчете определенных показателей в соответствии с предписанными для того или иного метода алгоритмами. В конечном результате высчитывается определенная числовая величина, которую сравнивают с табличными пороговыми значениями. Критерием достоверности будет результат сравнения полученной величины и табличного значения при данном числе наблюдений (или степеней свободы) и при заданном уровне безошибочного прогноза. Таким образом, в статистической процедуре оценки основное значение имеет полученный критерий достоверности, поэтому сам способ оценки достоверности в целом иногда называют тем или иным критерием по фамилии автора, предложившего его в качестве основы метода.

Оценка достоверности полученных статистических данных необходима, прежде всего, для определения **ошибки репрезентативности** – величины, показывающей насколько показатели, полученные при исследовании выборочной совокупности, отличаются от показателей, полученных при исследовании генеральной совокупности.

**Оценка достоверности результатов исследования с помощью параметрических методов.**

1. Одним из методов является определение средней ошибки средних величин. Она определяется для среднего арифметического:

$m = a / n$ , где  $m$  – ошибка среднего арифметического,  $a$  – среднее квадратичное отклонение,  $n$  – число наблюдений.

Ошибка относительного показателя определяется по формуле:

$$m \equiv \sqrt{\frac{p \times q}{n-1}},$$

где  $m$  – ошибка относительного показателя,  $p$  – показатель, выраженный в %, ‰, ‱.  $q$  – показатель, равный  $100 - p$  или  $1000 - p$  и т.д.,  $n$  – количество наблюдений. При  $n$  более 30 можно в формуле  $n-1$  заменить на  $n$  и полученной ошибкой пренебречь.

Результат считается достоверным ( $P$  или  $M$ ), если он, соответственно, превышает удвоенную или утроенную ошибку репрезентативности:  $M > 2-3 m$ ;  $P > 2-3 m$  (при  $n > 30$ ).

## 2. Определение доверительных границ средних и относительных величин

Формулы определения доверительных границ представлены следующим образом:

- для средних величин ( $M$ ):  $M_{\text{ген}} = M_{\text{выб}} \pm t * m_{\text{ср}}$
- для относительных показателей ( $P$ ):  $P_{\text{ген}} = P_{\text{выб}} \pm t * m_{\text{отн}}$

где  $M_{\text{ген}}$  и  $P_{\text{ген}}$  — соответственно, значения средней величины и относительного показателя генеральной совокупности;  $M_{\text{выб}}$  и  $P_{\text{выб}}$  значения средней величины и относительного показателя выборочной совокупности;  $m$  — ошибка репрезентативности;  $t$  — критерий достоверности (доверительный коэффициент).

Данный способ применяется в тех случаях, когда по результатам выборочной совокупности необходимо судить о размерах изучаемого явления (или признака) в генеральной совокупности.

Обязательным условием для применения способа является репрезентативность выборочной совокупности. Для переноса результатов, полученных при выборочных исследованиях, на генеральную совокупность необходима степень вероятности безошибочного прогноза ( $P$ ), показывающая, в каком проценте случаев результаты выборочных исследований по изучаемому

признаку (явлению) будут иметь место в генеральной совокупности.

При определении доверительных границ средней величины или относительного показателя генеральной совокупности исследователь сам задает определенную (необходимую) степень вероятности безошибочного прогноза  $P$ .

Для большинства медико-биологических исследований считается достаточной степень вероятности безошибочного прогноза  $P=95,5\%$ , т.е. число случаев генеральной совокупности, в которых могут наблюдаться отклонения от закономерностей, установленных при выборочном исследовании, не будет превышать  $5\%$ . При ряде исследований, связанных, например, с применением высокотоксичных веществ, вакцин, оперативного лечения и т.п., в результате чего возможны тяжелые заболевания, осложнения, летальные исходы, применяется степень вероятности  $P=99,7\%$ , т.е. не более чем у  $1\%$  случаев генеральной совокупности возможны отклонения от закономерностей, установленных в выборочной совокупности.

Заданной степени вероятности  $P$  безошибочного прогноза соответствует определенное, подставляемое в формулу, значение критерия  $t$ , зависящее также и от числа наблюдений.

*При  $n > 30$  степени вероятности безошибочного прогноза  $P=99,7\%$  соответствует значение  $t=3$ , а при  $P=95,5\%$  — значение  $t=2$ , при  $P=68\%$  — значение  $t=1$ . При  $n < 30$  величина  $t$  при соответствующей степени вероятности безошибочного прогноза определяется по специальной таблице (Н.А.Плохинского).*

### Значение критерия t по Плохинскому Н.А.

n (число наблюдений - 1)	(P) вероятность безошибочного прогноза			n (число наблюдений - 1)	(P) вероятность безошибочного прогноза		
	95%	99%	99,9%		95%	99%	99,9%
1	12,7	63,7	637,0	11	2,2	3,1	4,4
2	4,3	9,9	31,6	12	2,2	3,1	4,3
3	3,2	5,8	12,9	13	2,2	3,0	4,1
4	2,8	4,6	8,6	14-15	2,1	3,0	4,1
5	2,6	4,0	6,9	16-17	2,1	2,9	4,0
6	2,4	3,7	6,0	18-20	2,1	2,9	3,9
7	2,4	3,5	5,3	21-24	2,1	2,8	3,8
8	2,3	3,4	5,0	25-28	2,1	2,8	3,7
9	2,3	3,3	4,8	29-30	2,0	2,8	3,7
10	2,2	3,2	4,6				

n — число наблюдений минус единица (n — 1), P — вероятность безошибочного прогноза

### 3. Оценка достоверности разности результатов исследования (определение критерия t Стьюдента)

Данный способ применяется в тех случаях, когда необходимо определить, случайны или достоверны (существенны) различия между двумя средними величинами или относительными показателями, т.е. обусловлены ли эти различия каким-либо фактором или они случайны.

Обязательным условием для применения данного способа является репрезентативность выборочных совокупностей, а также предположение о наличии причинно-следственной связи разницы между сравниваемыми величинами (показателями) и факторами, влияющими на них.

Формулы определения достоверности разности представлены следующим образом:

$$t = \frac{M1 - M2}{\sqrt{m_1^2 + m_2^2}} - \text{для средних величин.}$$

$$t = \frac{P1 - P2}{\sqrt{m_1^2 + m_2^2}} - \text{для относительных величин.}$$



Если вычисленный критерий  $t$  более или равен 2 ( $t \geq 2$ ), что соответствует вероятности безошибочного прогноза  $P$ , равной или более 95,5% ( $P \geq 95,5\%$ ), то разность следует считать достоверной (существенной), т.е. обусловленной влиянием какого-то фактора, что будет иметь место и в генеральной совокупности.

При  $t < 2$  вероятность безошибочного прогноза  $P < 95,5\%$ . Это означает, что разность недостоверна, случайна, т.е. не обусловлена какой-то закономерностью (влиянием какого-то фактора).

*Поэтому полученный критерий должен всегда оцениваться по отношению к конкретной цели исследования.*

**Типичные ошибки, допускаемые исследователями при применении способа оценки достоверности разности результатов исследования:**

- При оценке достоверности разности результатов исследования по

критерию  $t$  часто делается вывод о достоверности (или недостоверности) самих результатов исследования. В действительности же этот способ позволяет судить только о достоверности (существенности) или случайности различий между результатами исследования.

- При полученном значении критерия  $t < 2$  часто делается вывод

о необходимости увеличения числа наблюдений. Если же выборочные совокупности репрезентативны, то нельзя делать вывод о необходимости увеличения числа наблюдений, так как в данном случае

значение критерия  $t < 2$  свидетельствует о случайности, недостоверности различия между двумя сравниваемыми результатами исследования.

**Применение непараметрических методов оценки достоверности полученных результатов**

**Непараметрические критерии оценки** — это совокупность статистических методов, которые позволяют оценить результаты исследований без вычисления общепринятых параметров.

Достоинства непараметрических методов (критериев) заключаются в том, что они не требуют знания характера распределения, могут применяться при любых распределениях, могут быть использованы при любом, даже небольшом числе наблюдений, применимы для признаков, имеющих количественное выражение, и признаков полуколичественного характера (например, степень тяжести и заболевания, результаты лечения и др.), относительно просты и не требуют проведения сложных расчетов, соответственно, экономят время при вычислении. Кроме того, непараметрические критерии обладают достаточной мощностью (чувствительностью).

В основе расчета непараметрических критериев лежит **упорядочивание (ранжирование)** имеющихся значений по отношению друг к другу, типа «больше-меньше» или «лучше-хуже». Это разграничение значений не предполагает точных количественных соотношений, а, следовательно, и ограничений на параметры и вид распределения. Поэтому для использования непараметрических критериев нужно меньше информации, нежели для критериев параметрических. В качестве оценок при непараметрических методах используются относительные характеристики — ранги, серии, знаки и др. Если в ситуации возможно применение параметрических критериев (нормальное распределение признака и незначительно различающееся разнообразие признака в совокупности), то им, как учитывающим большее количество информации, следует отдать предпочтение, так как они оказываются более мощными, чем непараметрические критерии, хотя и более трудоемкими.

Использование непараметрических критериев связано с такими понятиями, как «нулевая гипотеза» ( $H_0$ ), уровень значимости, достоверность статистических различий.

**«Нулевая гипотеза»** — это предположение о том, что в сравниваемых группах отсутствует различие в распределении частот.

**Уровень значимости** — это такая вероятность, которую принимают за основу при статистической оценке гипотезы. В качестве максимального уровня значимости, при котором нулевая гипотеза еще отклоняется, принимается 5%. При уровне значимости более 5% «нулевая гипотеза» принимается, различия

между сравниваемыми совокупностями принимаются статистически недостоверными, незначимыми.

Особого внимания заслуживает вопрос о мощности (чувствительности) критериев. Каждый из изучаемых критериев имеет характерную для себя мощность. Оценку значимости различий необходимо начинать с наименее мощного критерия. Если этот критерий опровергает нулевую гипотезу, то на этом анализ заканчивается. Если же нулевая гипотеза этим критерием не опровергается, то следует проверить изучаемую гипотезу более мощным критерием. Однако если значение характеристики, вычисленной для менее мощного критерия, оказалось очень далеким от критического значения, то мало надежды, что более мощный критерий опровергнет нулевую гипотезу.

### **Выбор непараметрических критериев для оценки результатов медицинских исследований (для определения существенности различий двух совокупностей).**

Для выбора того или иного критерия необходимо определить следующие моменты:

1. В каком виде получены результаты: в количественном или альтернативном (атрибутивном), т.е. представлены числом или альтернативной (атрибутивной, двухвариантной) оценкой: «есть признак» - «нет признака», «есть симптом» — «нет симптома» и т.д.

2. Связаны ли между собой сравниваемые выборочные совокупности или они взаимно независимы.

- *К связанным между собой* относятся выборочные совокупности с попарно сопряженными вариантами, например, при изучении количества гемоглобина в крови одних и тех же больных до и после лечения, различных физиологических показателей у спортсменов в норме, перед стартом и после окончания соревнований и т.п.

- *Взаимно независимые* совокупности не связаны между собой и могут иметь различную численность, например, результаты исследования крови у нескольких групп больных с различными стадиями болезни, результаты наблюдений над подопытной и контрольной группами исследования и т.д.

3. Сравниваются две или несколько выборочных совокупностей.

## **Применение непараметрических критериев для оценки различия 2 сопряженных (связанных) совокупностей**

Применяя критерии различия к связанным между собой выборочным совокупностям, исследователь стремится устранить или ослабить влияние на результаты оценки индивидуальной колеблемости вариант в пределах каждой совокупности, фиксируя основное внимание на изменениях каждого признака в динамике (например, до и после какого-либо воздействия). С этой целью, в частности, может быть использован и критерий Стьюдента — способ оценки достоверности разности результатов исследования. Однако применение этого способа связано с трудоемкой вычислительной работой.

Непараметрические критерии дают практически такую же информацию, но требуют для своего применения гораздо меньше вычислений. При этом более простые критерии (критерий знаков, максимум — критерий) обладают меньшей статистической мощностью; некоторое усложнение критерия (критерий Вилкоксона) приводит к повышению его мощности.

**Критерий знаков** в отличие от критерия  $t$  при оценке парных наблюдений (например, до и после лечения) учитывает не величину произошедших изменений, а только их направленность. Поэтому характер этих изменений учитывается в альтернативной форме (увеличение - уменьшение, ухудшение — улучшение и т.д., что для краткости обычно обозначается знаками «+» и «-», откуда и произошло название критерия). Случаи, когда парные наблюдения не имеют разницы (что можно обозначить знаком = или 0), из дальнейшего сравнения исключаются. В связи с этим следует стремиться, чтобы количество таких нулевых разностей было минимальным (обеспечение непрерывности выборочных данных путем повышения точности измерения количественных и полуколичественных наблюдений).

Если число положительных изменений близко к числу отрицательных изменений, то очевидно, что различия между сравниваемыми выборочными совокупностями не могут быть признаны статистическими значимыми. Наоборот, вероятность значимого различия возрастает в случаях заметной направленности изменений в одну из сторон, т.е. в случаях преобладания одного из знаков.

**Максимум – критерий.** Это более мощный критерий, основанный уже на величине происшедших изменений. Для этого:

- 1) определяют разности в парах наблюдений с учетом знаков;
- 2) располагают разности по их абсолютным величинам;
- 3) определяют число первых наибольших разностей с одинаковым знаком, т.е. до величины с противоположным направлением изменения.

Оценка ведется по стандартным значениям: 6 пар наблюдений с одним знаком — 5% риска ошибки; 8 пар наблюдений — 1 % риска ошибиться в достоверности различий и 11 пар наблюдений — менее 1% риска ошибки.

**Критерий Вилкоксона.** До настоящего времени при оценке различий двух связанных совокупностей рассматривалось лишь направление оказания действия и в какой-то степени величина разностей в паре наблюдений. Для более точного суждения о достоверности различий принимаются во внимание размеры этих разностей.

Вычисление критерия Вилкоксона осуществляется в следующей последовательности:

1. Вычисляются разности в парах наблюдений.
2. Проставляются ранги по величине разности без учета знаков (от меньшей разности к большей, результаты без изменений исключаются).
3. Подсчитывается сумма однозначных рангов.
4. Оценивается меньшая из сумм.

### **Применение непараметрических критериев для оценки достоверности показателей исследований для несвязанных совокупностей**

Эти критерии особенно часто применяются в исследованиях, где имеются опытные и контрольные группы, где необходимо сравнить результаты двух групп наблюдений, относящихся к различным заболеваниям или стадиям болезни и т.д.

#### **Критерий Уайта**

Последовательность расчета:

1. Данные рядов X и Y ранжируются от меньшей величины к большей вне зависимости от их принадлежности к тому или иному ряду.
2. Ранги суммируются отдельно для рядов X и Y.

3.Меньшая из сумм оценивается по таблице.

### **Критерий Колмогорова-Смирнова**

Это наиболее мощный критерий из серии непараметрических критериев, применяемых при сопоставлении двух различных групп наблюдений. Задача его та же, что и всех предыдущих.

### **Применение непараметрических критериев для определения существенности различий любых совокупностей**

#### **Критерий соответствия «хи-квадрат»**

С помощью хи-квадрат определяют соответствие (согласие) эмпирического распределения теоретическому, и тем самым оценивают достоверность различий между выборочными совокупностями. Критерий применяется в тех случаях, когда нет необходимости знать величину того или иного параметра (среднюю или относительную величину) и требуется оценить достоверность различий не только двух, но и большего числа групп.

Критерий соответствия применяется для статистической оценки результатов исследования в случаях, когда нет необходимости знать величину самого показателя, размер связи, а требуется лишь подтвердить, существенно ли влияние изучаемого фактора или оно случайно, и подтвердить наличие взаимосвязи между явлениями. В отличие от метода оценки достоверности по критерию Стьюдента, который позволяет проводить только попарное сравнение, критерий соответствия применяется для сопоставления не только двух, но и большего числа групп, в этом его преимущество. Определение критерия соответствия основано на довольно распространенном в исследованиях приеме: доказывать от противного.

В практическом здравоохранении метод  $\chi^2$  может широко использоваться при оценке эффективности прививок, действия препаратов, результатов различных методов лечения и профилактики заболеваний влияния условий труда и быта на заболеваемость работающих. С помощью критерия можно определить, влияют или нет сроки госпитализации на течение заболеваемости, влияет ли материальное обеспечение населения на уровень заболеваемости и т.д.

Критерий определяется по формуле:

$$\chi^2 = \sum \frac{(P - P_1)^2}{P_1},$$

где  $P$  — фактические (эмпирические) данные;  $P_1$  — «ожидаемые» (теоретические) данные, вычисленные на основании нулевой гипотезы.

Определение критерия основано на расчете разницы между фактическими и «ожидаемыми» данными. Чем больше эта разность  $(P - P_1)$ , тем с большей вероятностью можно утверждать, что существуют различия в распределении сравниваемых выборочных совокупностей, и наоборот, чем меньше разность, тем меньше шансов на то, что сравниваемые выборочные совокупности различны между собой.

### Задача – эталон №1

**Условие.** Исследована заболеваемость гриппом у школьников. Из 1200 школьников 800 привиты от гриппа. Среди привитых заболеваемость гриппом 8%, среди не привитых – 15%. Оцените эффективность вакцинации.

**Решение.** Для начала определим ошибку относительных величин:

$$m_{\text{непривитых}} = \sqrt{\frac{p \times q}{n-1}} = \sqrt{\frac{15 \times 85}{400}} = 0,96$$

$$m_{\text{привитых}} = \sqrt{\frac{p \times q}{n-1}} = \sqrt{\frac{8 \times 92}{800}} = 1,78$$

Определим теперь достоверность разницы 2 относительных величин:

$$t = \frac{P_1 - P_2}{\sqrt{m_1^2 + m_2^2}} = \frac{15 - 8}{\sqrt{0,96^2 + 1,78^2}} = 3,5$$

Получив такие данные, можем говорить о том, что вакцинация эффективна с вероятностью 99%, т.к.  $t \geq 3$ .

### Задача – эталон №2

**Условие.** Средняя длительность лечения больных пневмонией с применением препарата А составила  $20,5 \pm 0,8$  дней, а при применении препарата В –  $22,5 \pm 0,7$  дня. Оцените

эффективность препаратов. Какой из препаратов более эффективен?

**Решение.** Имеет средние величины  $M_1=20,5$ ,  $M_2=22,5$ , и значения ошибок средних величин  $m_1=0,8$  и  $m_2=0,7$ . Подставим эти значения в формулу для определения критерия Стьюдента и рассчитаем его:

$$t = \frac{M_1 - M_2}{\sqrt{m_1^2 + m_2^2}} = \frac{20,5 - 22,5}{\sqrt{0,8^2 + 0,7^2}} = 0,5$$

Получили критерий Стьюдента  $t = 0,5 \leq 2$ , т.е. достоверной разницы между препаратами нет и лечение препаратом А и В одинаково эффективно.

### Задача – эталон №3

**Условие.** Определить отличаются ли сроки постановки диагноза больным ревматизмом в поликлинике №1 (принимает кардиолог) и в поликлинике №2 (принимает терапевт), если имеются такие данные:

Прием ведет:	Срок $\leq 15$ дней	Срок $\geq 15$ дней	Всего
Кардиолог	54	19	73
Терапевт	7	14	21
<b>Всего</b>	<b>61</b>	<b>33</b>	<b>94</b>

**Решение.** Определяем ожидаемые величины на основании «нулевой гипотезы». Т.е. предполагаем, что независимо от того, кто ведет прием, сроки постановки диагноза одинаковы. В этом случае распределение 2 групп больных, обслуживаемых с участием кардиолога и терапевта, должно быть одинаково и соответствовать итоговому фактическому распределению пациентов, т.е. 61 и 33. При таком условии в 1 группе (кардиолог) количество пациентов с установленным диагнозом ранее 15 дней определяется по пропорции:

$$\begin{aligned} &94 - 61 \\ &73 - x \\ &x = (61 \cdot 73) / 94 = 47,4 \end{aligned}$$

Соответственно, число больных с установленным диагнозом позднее 15 дней =  $73 - 47,4 = 25,6$ . Аналогично поступаем с данными по поликлинике, в которой прием ведет терапевт.

Далее определяем разницу между фактическими и ожидаемыми числами  $(P-P_1)$  и затем рассчитываем квадрат этой разности  $(P-P_1)^2$ .

Все полученные данные заносим в таблицу.

Прием	Число больных	Р		Р <sub>1</sub>		Р-Р <sub>1</sub>		(Р-Р <sub>1</sub> ) <sup>2</sup>		(Р-Р <sub>1</sub> ) <sup>2</sup> /Р	
		$\leq 15$	$\geq 15$	$\leq 15$	$\geq 15$	$\leq 15$	$\geq 15$	$\leq 15$	$\geq 15$	$\leq 15$	$\geq 15$
Кардиолог	73	54	19	47,5	25,6	+6,6	-6,6	43,6	43,6	0,9	1,7



Терапевт	21	7	14	13,6	7,4	-6,6	+6,6	43,6	43,6	3,2	5,9
Всего	94	61	33								

Критерий хи-квадрат рассчитываем по формуле:

$$\chi^2 = \sum \frac{(P - P_1)^2}{P_1} = 0,9 + 3,2 + 1,7 + 5,9 = 11,7$$

Т.е. полученный критерий сильно отличается от 0, что позволяет думать об ошибочности нашей «нулевой» гипотезы.

### **Задачи для самостоятельной работы**

**1.** В результате лечения 263 больных туберкулезом легких с применением пневмоторакса рецидивы отмечались у 19 человек, в то же время при лечении сочетанным способом 77 больных рецидивы наблюдались у 15 человек. Докажите эффективность сочетанной терапии.

**2.** При лечении бронхита экспериментальным препаратом А было выявлено 25% осложнений из 74 случаев. При использовании стандартной схемы лечения – 39% из 119 случаев. Оцените эффективность экспериментального препарата.

### **Контрольные вопросы**

1. Цели проведения оценки вероятности результатов исследования
2. Методы оценки вероятности результатов исследования
3. Этапы оценки вероятности результатов статистического исследования
4. Определение ошибки репрезентативности
5. Средняя ошибка средней арифметической и относительной величины
6. Доверительные интервалы
7. Методика определения доверительных интервалов
8. Критерий Стьюдента. Применение. Формулы. Определение. Значение
9. Примеры использования непараметрических методов оценки вероятности результатов исследования в практической медицине

## Метод стандартизации

При изучении общественного здоровья и здравоохранения в научных или практических целях исследователю нередко приходится доказывать влияние факторных признаков на результативные при сравнении двух или более совокупностей. С этой целью применяется целый ряд статистических приемов.

При сравнении двух неоднородных совокупностей по какому-либо признаку (составу) применяются методы стандартизации (прямой, обратный, косвенный).

В данном учебном пособии рассматривается прямой метод стандартизации. Этот метод применяется при наличии полных сведений, как о составе сравниваемых совокупностей, так и распределении в них явления.

**Условие применения метода стандартизации.** Метод применяется при сравнении интенсивных показателей в совокупностях, отличающихся по составу (например, по возрасту, полу, профессиям и т.д.).

**Сущность метода стандартизации.** Он позволяет устранить (элиминировать) возможное влияние различий в составе совокупностей по какому-либо признаку на величину сравниваемых интенсивных показателей. С этой целью составы совокупностей по данному признаку уравниваются, что в дальнейшем позволяет рассчитать стандартизованные показатели.

**Стандартизованные показатели** — условные, гипотетические величины, они не отражают истинных размеров явлений. Стандартизованные показатели свидетельствуют о том, каковы были бы значения сравниваемых интенсивных показателей, если бы были исключены различия в составах совокупностей.

**Назначение метода стандартизации.** Таким образом, метод стандартизации применяется для выявления влияния фактора неоднородности составов совокупностей по какому-либо признаку на различия сравниваемых интенсивных показателей.

### Этапы расчета стандартизованных показателей

**I этап.** Расчет общих и частных интенсивных показателей: общих — по совокупностям в целом; частных — по признаку различия (по полу, возрасту, стажу работы и т.д.).

**II этап.** Определение стандарта, т.е. выбор одинакового численного состава среды по данному признаку (по возрасту, полу и т.д.) для сравниваемых совокупностей. Как правило, за стандарт принимается сумма или полусумма численностей составов соответствующих групп. В то же время стандартом может стать состав любой из сравниваемых совокупностей, а также состав по аналогичному признаку какой-либо другой совокупности. Например, при сравнении летальности в конкретной больнице по двум отделениям скорой помощи за стандарт может быть выбран состав больных любой другой больницы скорой помощи. Таким образом, так или иначе уравниваются условия среды, что дает возможность провести расчеты новых чисел явления, называемых «ожидаемыми величинами».

**III этап.** Вычисление ожидаемых абсолютных величин явления в группах стандарта на основе групповых интенсивных показателей, рассчитанных на I этапе. Итоговые числа по сравниваемым совокупностям являются суммой ожидаемых величин в группах.

**IV этап.** Вычисление стандартизованных показателей для сравниваемых совокупностей, используя итоговые ожидаемые величины в группах и новую среду - стандарт.

**V этап.** Сопоставление соотношений стандартизованных и интенсивных показателей, формулировка вывода.

### **Задача – эталон №1**

**Условие.** Число осложнений у больных в стационарах А и В измерена в баллах. Результаты приведены ниже в виде таблицы. Оцените качество медицинской помощи в стационарах А и В на основании количества возникающих осложнений.

Тяжесть осложнений	Стационар А		Стационар В	
	Число больных	Число осложнений	Число больных	Число осложнений
До 10	250	20	300	22
11-20	450	42	450	41
21- 30	120	22	250	45
31-40	85	25	220	60
Более 40	30	15	100	44

**Решение.** Стоим рабочую таблицу. В столбцах 1,2,3,4,5 приведены для удобства исходные данные. Начнем расчет стандартизированных показателей.

В столбце 6 приведена относительная частота осложнений в стационаре А. Рассчитывается как  $3/2 * 100\%$ , где 3 и 2 номер столбцов.

В столбце 7 приведены те же данные, что и в столбце 6, только для стационара В. Рассчитываются как  $5/4 * 100\%$ , где 5 и 4 номера столбцов.

Далее мы принимаем правильное утверждение о том, что количество осложнений в стационарах А и В одинаково. В столбце 10 приведем сумму всех больных в обоих стационарах. Таким образом, получим стандартизированные показатели по количеству поступивших больных для обоих стационаров. Рассчитывается как  $2+4$ , где 2 и 4 номера столбцов.

Тяжесть осложнений	Стационар А		Стационар В		Относительные показатели осложнений		Ожидаемое число больных с осложнениями		Стандарты
	Больные	Осложнения	Больные	Осложнения			Стационар А	Стационар В	
					Стационар А	Стационар В			
1	2	3	4	5	6	7	8	9	10
≤10	250	20	300	22	8	7,3	44	40,1	550
11-20	450	42	450	41	9,3	9,1	83,7	81,9	900
21-30	120	22	250	45	18,3	18	67,7	66,6	370
31-40	85	25	220	60	29,4	27,3	89,7	83,3	305
≥ 40	30	15	100	44	50	44	65	57,6	130
Всего	935	124	1310	212	13,3	16,1	351,6	329,1	2245

В столбце 8 рассчитаем стандартизированное предполагаемое количество больных, у которых возникнут осложнения в

стационаре А, исходя из знания стандартизированного количества больных (10) и знания относительного показателя осложнений для стационара А (6). Рассчитывается как  $6 * 10 / 100\%$ , где 6 и 10 номера столбцов. Также, только для стационара В, проведем расчеты в столбце 9.

Данные «Всего» в столбцах 8 и 9 получим, суммируя все данные по соответствующим столбцам (8 и 9 соответственно).

Данные строки «Всего» в столбцах 8 и 9 являются стандартизированными показателями количества осложнений в стационарах А и В соответственно. Для большей наглядности выразим их в виде относительной величины интенсивного показателя по отношению к стандартизированному показателю количества больных (столбец 10, строка «Всего»):

Стационар	Интенсивный показатель по отношению к числу больных в каждом из стационаров	Стандартизированный показатель по отношению к стандартизированному количеству больных
А	13,3%	$\frac{351,6}{2245} \times 100\% = 15,6\%$
В	16,1%	$\frac{329,1}{2245} \times 100\% = 14,6\%$

**Вывод.** Полученные стандартизированные показатели говорят о более высокой частоте осложнений в стационаре А (15,6%) и более низкой в стационаре В (14,6%). При анализе осложнений в стационаре А и В, учитывая неоднородность состава больных, сравнение показателей частоты осложнений возможно только после исключения неоднородности в тяжести состояния больных. Рассчитанные стандартизированные показатели говорят о том, что при абсолютно однородном составе больных в обоих стационарах, число осложнений в стационаре А было бы на 1% больше, чем в стационаре В.

### Задачи для самостоятельного решения

**1.** Среди работающих на двух предприятиях ("Явор" и "АВЕС"), которые не отличаются производственными условиями, но отличаются структурой распределения по полу, необходимо провести анализ распространенности язвенной болезни двенадцатиперстной кишки. Исходные данные представленные в таблицы.

Пол	Предприятие "Явор"		Предприятие "АВЕС"	
	К-во работающих	К-во больных	К-во работающих	К-во больных
Мужской	400	8	200	4

Женский	200	6	400	12
Всего	600	14	600	16

**2.** Среди работающих на двух предприятиях ("1" и "2"), которые не отличаются производственными условиями, но отличаются структурой распределения по полу необходимо провести анализ распространенности травматизма. Исходные данные представлены в таблице.

Пол	Предприятие "1"		Предприятие "2"	
	Кол-во работающих	Кол-во травмированных	Кол-во работающих	Кол-во травмированных
Мужской	300	8	200	4
Женский	300	6	400	12
Всего	600	14	600	16

**3.** В чем состоит практическое значение метода стандартизации? Дайте полный ответ и обоснуйте его.

**А.** Позволяет оценить динамику показателей

**В.** Исключает влияние неоднородности составов исследуемых групп

**С.** Позволяет сравнить частоту явлений, изучаемых в неоднородных группах

**Д.** Позволяет оценить влияние исследуемого фактора на величину общих показателей

**Е.** Позволяет сравнивать интенсивные показатели

### **Контрольные вопросы**

1. Для чего используется метод стандартизации?
2. Какие существуют методы стандартизации?
3. В чем суть прямого метода стандартизации?
4. Назовите этапы прямого метода стандартизации
5. Какие критерии и методы отбора стандарта?
6. О чем свидетельствуют стандартизированные показатели?
7. Назовите случаи использования метода стандартизации в практике врача

## **Корреляционно – регрессионный метод в оценке медико-биологической информации**

При изучении общественного здоровья и здравоохранения в научных и практических целях исследователю часто приходится проводить статистический анализ связей между факторными и результативными признаками статистической совокупности (причинно-следственная связь) или определение зависимости параллельных изменений нескольких признаков этой совокупности от какой-либо третьей величины (от общей их причины). Необходимо уметь изучать, особенности этой связи, определять ее размеры и направление, а также оценивать ее достоверность. Для этого используются методы корреляции.

**Функциональная связь** — такой вид соотношения между двумя признаками, когда каждому значению одного из них соответствует строго определенное значение другого (площадь круга зависит от радиуса круга и т.д.). Функциональная связь характерна для физико-математических и химических процессов. Коэффициент корреляции для такого вида связи равен  $\pm 1$ .

**Корреляционная связь** — такая связь, при которой каждому определенному значению одного признака соответствует несколько значений другого взаимосвязанного с ним признака (связь между ростом и массой тела человека; связь между температурой тела и частотой пульса и др.). Корреляционная связь характерна для социально-гигиенических процессов, клинической медицины и биологии.

### **Практическое значение корреляционной связи:**

- Выявление причинно-следственной связи между факторными и результативными признаками (при оценке физического развития, для определения связи между условиями труда, быта и состоянием здоровья, при определении зависимости частоты случаев болезни от возраста, стажа, наличия производственных вредностей и др.).
- Зависимость параллельных изменений нескольких признаков от какой-то третьей величины. Например, под воздействием высокой температуры в цехе происходят изменения кровяного давления, вязкости крови, частоты пульса и др.

Определение характера корреляционной связи производится методом расчета коэффициента корреляции.

**Коэффициент корреляции** - одним числом дает представление о направлении и силе связи между признаками (явлениями); пределы его колебаний от 0 до  $\pm 1$ .

**Корреляционная связь может быть:**

- **по направлению:**

- а) прямая (положительный коэффициент корреляции)
- б) обратная (отрицательный коэффициент корреляции)

- **по силе:**

- а) сильная: коэффициент корреляции равен  $\pm 0,7$  до  $\pm 1$ ;
- б) средняя: коэффициент корреляции равен  $\pm 0,3$  до  $\pm 0,699$ ;
- в) слабая: коэффициент корреляции равен 0 до  $\pm 0,299$ .

- **по достоверности:**

- а) достоверная;
- б) недостоверная.

- **по форме:**

- а) линейная (увеличение значения 1 признака в  $n$  раз приводит к увеличению второго признака в такое же количество раз);
- б) нелинейная (нет линейной зависимости между признаками, график представляет собой кривую).

**Виды коэффициента корреляции:**

1. **Коэффициент парной корреляции** – дает представление о влиянии и связи между 2 явлениями без учета влияния посторонних факторов. Выделяют:

- Ранговый (Спирмена)
- Линейный (Пирсона)

2. **Множественный коэффициент корреляции** – показывает связь между 2 и более признаками. Его определение позволяет иметь представление о взаимодействии многих факторов и их влиянии на основное изучаемое явление.

3. **Парциальный коэффициент корреляции** – оценивается на основании парного и множественного коэффициента корреляции и позволяет определить взаимосвязь между 2 факторами, исключая воздействие других факторов.

**Ранговый коэффициент корреляции Спирмена (ранговый, непараметрический метод)**



Метод ранговой корреляции Спирмена позволяет определить силу и направление корреляционной связи между двумя признаками или двумя иерархиями признаков.

Для подсчета ранговой корреляции необходимо располагать двумя рядами значений, которые могут быть проранжированы. Такими рядами могут быть:

**А)** Два признака, измеренные в одной и той же группе переменных (наиболее часто в этом качестве выступает группа людей, которых принято тогда именовать испытуемыми или респондентами. Естественно, под переменными подразумеваются не сами люди, а данные ими ответы на те или иные вопросы).

**Б)** Две индивидуальные иерархии признаков, выявленные у двух испытуемых по одному и тому же набору признаков (скажем, по ответам на пункты анкеты или теста).

**В)** Две групповые иерархии признаков (например, соответствие каких-либо выборов, сделанных одной группой людей выборам другой группы).

**Г)** Индивидуальная и групповая иерархии признаков (например, сопоставление индивидуальной иерархии жизненных ценностей сотрудника усредненному мнению группы на этот же счет; сопоставление последовательности товаров, которые приобрели бы (в среднем) жители города А и города Б при условии получения премии, на которую заранее не рассчитывали.)

**Ограничения метода ранговой корреляции.** По каждой переменной должно быть представлено не менее 5 наблюдений. Верхняя граница выборки – меньше или равна 40. Коэффициент ранговой корреляции Спирмена при большом количестве одинаковых рангов по одной или обоим сопоставляемым переменным дает огрубленные значения. В идеале оба коррелируемых ряда должны представлять собой две последовательности несовпадающих значений. В случае несоблюдения такого условия вносится поправка на одинаковые ранги. Помимо этих ограничений, следует так же помнить об ограничениях корреляционного метода вообще – невозможность обнаружения причинной связи между явлениями.

**Коэффициент корреляции Спирмена вычисляется по формуле:**

$$r = \frac{6 \times \sum d^2}{n \times (n^2 - 1)},$$

где  $d$  – разность рангов (ранг положение данного значения в таблице по возрастанию),  $n$  – количество значений.

### **Линейный коэффициент корреляции Пирсона**

Предположим, что мы располагаем выборкой данных о какой-то группе объектов. Пусть эти объекты обладают общими родовыми особенностями (примерно одинаковы). Пусть, к тому же, у каждого из объектов можно количественно измерить, как минимум, два каких-либо параметра. При этих обстоятельствах открывается возможность для подсчета линейной корреляции между двумя (или более) признаками, присущими этим объектам.

Например, такими выборками данных могут служить сведения о:

- группе людей, рост и вес тела которых мы измеряем;
- длине и ширине лепестка какого-нибудь цветка;
- длине ствола оружия и начальной скорости пули;
- величине IQ и времени решения учебной задачи;
- длине тела и длине хвоста крокодила или тигра (если найдется желающий их измерять), и т.д.

Во всех этих примерах имеется возможность определить корреляцию, то есть – степень согласованности в изменении двух признаков. «Чем больше крокодил, тем длиннее ли его хвост?» «Решают ли люди с высоким коэффициентом интеллекта задачи такого-то типа быстрее, чем с низким и средним?»

**Требования и ограничения.** Необходимо иметь в виду, что сопоставляемые характеристики должны быть, во-первых, внутренне присущи объектам и, во-вторых, быть количественно измеряемыми. Ввиду того, что расчет линейной корреляции проводится с использованием средних значений и дисперсий, следует также помнить, что эта процедура относится к разряду параметрических методов и, соответственно, требует нормальности распределения признака.

Также следует помнить, что никакая корреляция вообще не устанавливает зависимости одного обстоятельства от другого, а лишь является мерой совместной вариации двух величин. И, наконец, линейная корреляция потому и называется линейной,

что способна дать ответ о взаимосвязи изменений того и иного свойства объекта только тогда, когда возрастание - убывание значения признака происходит по линейному закону (график – прямая линия).

**Коэффициент корреляции Пирсона вычисляется по формуле:**

$$r = \frac{\sum d_x \times d_y}{\sqrt{\sum d_x^2 \times d_y^2}},$$

где  $d_x$  и  $d_y$  – отклонение каждой варианты от среднего арифметического.

Для любого коэффициента корреляции резонно вычислять **ошибку коэффициента корреляции:**

$$m = \pm \sqrt{\frac{1 - p^2}{n - 2}}, \quad m = \pm \sqrt{\frac{1 - r^2}{n - 2}}$$

где  $m$  – ошибка коэффициента корреляции,  $p$  – ранговый коэффициент корреляции,  $r$  – линейный коэффициент корреляции,  $n$  – число наблюдений.

Для оценки ошибки коэффициента корреляции необходимо использовать формулу:

$$t = \frac{p}{m} = \frac{r}{m} \geq 3,$$

т.е. полученная ошибка коэффициента корреляции должна быть меньше самого коэффициента корреляции в 3 и более раз.

**Регрессионный анализ.** Регрессия – это функция, позволяющая вычисление среднего значения в выборке одного корреляционного признака по известной величине второго корреляционного признака (на сколько изменится значение одного признака при увеличении (уменьшении) второго на 1).

Для расчета используется формула:

$$R_{y/x} = r \times \frac{\delta_y}{\delta_x}$$

Полученный коэффициент регрессии подставляется в уравнение регрессии:

$$y = y_1 + R_{y/x} * (x - x_1)$$

**Задача – эталон №1**

**Условие.** Определить зависимость между частотой бронхитов и длительностью курения, исходя из данных приведенной ниже таблицы.

<b>Длительность курения, лет</b>	<b>Частота бронхитов, %</b>
3	6
4	9
5	12
6	13
7	14
8	21
9	26
10	35

**Решение.** Видим классическую задачу на умения применять коэффициент корреляции – необходимо определить зависимость между двумя явлениями. Для выбора метода определения коэффициента корреляции обратим внимание на то, что нам нужно определить степень согласованности в изменении двух признаков. Такое условие подталкивает нас к решению задачи с использованием коэффициента Пирсона. Для удобства при расчетах составим рабочую таблицу:

Длительность курения, лет, x	Частота бронхитов, %, y	dx	dy	dx*dy	dx <sup>2</sup>	dy <sup>2</sup>
3	6	-3,5	-11	38,5	12,25	121
4	9	-2,5	-8	20	6,25	64
5	12	-1,5	-5	7,5	2,25	25
6	13	-0,5	-4	2	0,25	16
7	14	0,5	-3	-1,5	0,25	9
8	21	1,5	4	6	2,25	16
9	26	2,5	9	22,5	6,25	81
10	35	3,5	18	63	12,25	324
X <sub>среднее</sub> =6,5	Y <sub>среднее</sub> =17					

$$\sum dx \times dy = 158, \sum dx^2 = 42, \sum dy^2 = 656$$

Таким образом, имеем все данные для расчета линейного коэффициента корреляции Пирсона:

$$r = \frac{\sum d_x \times d_y}{\sqrt{\sum d_x^2 \times \sum d_y^2}} = \frac{158}{\sqrt{42 + 656}} = +0,97$$

Определим ошибку полученного коэффициента корреляции:

$$m = \pm \sqrt{\frac{1-r^2}{n-2}} = \sqrt{\frac{1-0,97^2}{8-2}} = \pm 0,1$$

**Вывод.** Исходя из проведенных расчетов по предоставленным результатам, можно сделать вывод, что связь между длительностью курения и частотой развития бронхитов сильная, прямая, достоверная.

### Задачи для самостоятельного решения

**1.** При изучении физического развития шестилетних мальчиков были полученные следующие данные: длина тела - 105, 119, 118, 107, 110, 116, 117, 118 (см), масса тела - 28, 33, 32, 29, 30, 31, 32, 30 (кг). Определите и оцените взаимосвязь между массой и длиной тела обследованных мальчиков.

**2.** Анализируя заболеваемость по отдельным районам, специальная комиссия областного управления здравоохранения констатировала, что план профилактических мероприятий среди сельскохозяйственных работников (в 10 районах) выполнен на: 52%, 57%, 64%, 68%, 72%, 75%, 79%, 83%, 86%, 88%. Частота же заболеваний кожи и подкожной клетчатки с временной утратой трудоспособности в этих же районах составляет (в расчета на 100 работающих): 3.2, 3.5, 2.9, 2.7, 2.4,

2.8, 2.9, 2.2, 1.7, 1.5. В какой мере частота этих случаев нетрудоспособности связана с выполнением профилактических мероприятий среди работников сельскохозяйственного производства?

**3.** Определить коэффициенты корреляции и регрессии по результатам изучения заболеваемости населения острыми кишечными инфекциями (на 10 тыс. населения) и состояния шахтных колодцев (по комплексной оценке) в некоторых населенных пунктах.

Показатель	Населенный пункт				
	1.	2.	3.	4.	5.
Заболеваемость	49,8	57,7	59,4	60,2	63,6
Удельный вес шахтных колодцев в неудовлетворительном состоянии	23	27	34	29	38

### Контрольные вопросы

1. Виды связей между явлениями
2. Классификация корреляционной связи
3. Виды коэффициента корреляции
4. Методы вычисления коэффициента корреляции
5. Область применения коэффициента Спирмена
6. Область применения коэффициента Пирсона
7. Как оценивается вероятность и достоверность коэффициента корреляции?
8. Что такое регрессия?
9. Методы вычисления коэффициента регрессии
10. Уравнение регрессии
11. Область применения корреляционно-регрессионного анализа в практической медицине

# Динамические ряды. Анализ рядов динамики.

## Медико-биологическое значение и область применения

**Динамический ряд** — это ряд однородных статистических величин, показывающих изменение явления во времени и расположенных в хронологическом порядке.

Динамический ряд может быть представлен *абсолютными числами* (изменение числа больных), *средними величинами* (среднее число лабораторных анализов за неделю) и *относительными показателями* (изменение рождаемости, заболеваемости, травматизма, обеспеченности врачами).

Числа, из которых состоит динамический ряд, называются **уровнями ряда**.

Классификация динамических рядов:

1. По простоте:

- **Простые** – одному временному показателю соответствует один количественный показатель. Например, заболеваемость гипертонической болезнью за период 1990 – 99 гг. на участке №22 поликлиники № 4 города N.

Год	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Число больных	32	34	36	51	67	12	25	67	88	77

- **Сложные** – одному временному показателю соответствует несколько показателей динамического ряда. Например, заболеваемость гипертонической болезнью за период 1990 – 99 гг. на участке №22 поликлиники №4 города N с учетом числа осложненных случаев.

Год	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Число больных	32	34	36	51	67	12	25	67	88	77
Число осложненных случаев	12	11	1	23	54	4	22	54	12	11

2. По временным характеристикам:

- **Моментные** – состоят из величин, характеризующих изучаемое явление на определенный момент времени. Уровни ряда не суммируются. Особенностью моментного ряда динамики является то, что в его уровни могут входить

одни и те же единицы изучаемой совокупности. Хотя и в моментном ряду есть интервалы – промежутки между соседними в ряду датами, - величина того или иного конкретного уровня не зависит от продолжительности периода между двумя датами. Посредством моментных рядов динамики в медицинской статистике изучаются запасы медикаментов в АПУ, состояние кадров, количество медицинского оборудования и других показателей, отображающих состояние изучаемых явлений на отдельные даты (моменты) времени. Примером моментного динамического ряда является количество работников поликлиники №4 города N по годам за период с 1990 по 1999 гг.

Год	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Число врачей	115	111	100	89	67	82	84	89	123	131

- **Интервальные** – состоят из величин, отображающих итоги за определенные интервалы времени. Каждый уровень интервального ряда уже представляет собой сумму уровней за более короткие промежутки времени. При этом единица совокупности, входящая в состав одного уровня, не входит в состав других уровней. Особенностью интервального ряда динамики является то, что каждый его уровень складывается из данных за более короткие интервалы (субпериоды) времени. Например, суммируя рождаемость за первые три месяца года, получают ее объем за I квартал, а суммируя рождаемость за четыре квартала, получают ее величину за год, и т. д. При прочих равных условиях уровень интервального ряда тем больше, чем больше длина интервала, к которому этот уровень относится. Свойство суммирования уровней за последовательные интервалы времени позволяет получить ряды динамики более укрупненных периодов. Посредством интервальных рядов динамики в медицинской статистике изучают показатели, отображающие итоги функционирования изучаемого явления за отдельные периоды. Например, количество больных ВИЧ-инфекцией в городе N по годам за период 1990 – 99 гг.

Год	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999
Число больных ВИЧ	2	19	168	211	356	812	1024	1997	2111	3561

**Анализ динамического (временного) ряда** сводится к вычислению следующих показателей:



- абсолютного прироста (или снижения)
- темпа роста (или снижения)
- темпа прироста
- значения 1% прироста

**Абсолютный прирост** представляет собой разность между последующим и предыдущим уровнем динамического ряда. Вычисляется по формуле:

$$АП = x_2 - x_1 = x_n - x_{n-1},$$

где АП – абсолютный прирост,  $x$  – значение данного уровня ряда,  $n$  – порядковый номер данного уровня ряда в ряду динамики. Имеет значение для определения абсолютного числа прироста в ряду динамики, например, определение годовой заболеваемости ВИЧ-инфекцией в городе N.

**Темп роста** — это отношение последующего уровня к предыдущему, умноженное на 100%. Вычисляется по формуле:

$$ТР = \frac{x_2}{x_1} \times 100\% = \frac{x_n}{x_{n-1}} \times 100\%,$$

где ТР – темп роста,  $x$  – значение данного уровня ряда,  $n$  – порядковый номер данного уровня ряда в ряду динамики. Имеет значение для определения прироста в ряду динамики, выраженного в процентах или других относительных величинах. Например, определение годовой заболеваемости ВИЧ-инфекцией в городе N, выраженное в процентах.

**Темп прироста** - отношение абсолютного прироста (снижения) к предыдущему уровню динамического ряда, умноженное на 100%. Вычисляется по формуле:

$$ТП = \frac{x_n - x_{n-1}}{x_{n-1}} \times 100\%,$$

где ТП – темп прироста,  $x$  – значение данного уровня ряда,  $n$  – порядковый номер данного уровня ряда в ряду динамики. Используется для наглядного отображения в процентах прибавления (уменьшения) какого-либо показателя за определенный период времени. Например, темпы прироста больных ВИЧ-инфекцией за каждый из анализируемых годов.

**Абсолютное значение 1% прироста** определяется отношением абсолютного прироста к темпу прироста. Вычисляется по формуле:

$$AZ = \frac{x_n - x_{n-1}}{ТП},$$

где АЗ – абсолютное значение 1% прироста, ТП – темп прироста,  $x$  – значение данного уровня ряда,  $n$  – порядковый номер данного уровня ряда в ряду динамики. Используется для получения абсолютных цифр прироста, при условии, что имеется данные о темпах прироста и рассматриваемый динамический ряд.

**Методы сглаживания и выравнивания динамических рядов.** Исключение случайных колебаний значений уровней ряда осуществляется с помощью нахождения “усредненных” значений.

**Способы устранения случайных факторов (сглаживания)** делятся на две больше группы:

1. Способы “механического” сглаживания колебаний путем усреднения значений ряда относительно других, расположенных рядом, уровней ряда.
2. Способы “аналитического” выравнивания, т. е. определения сначала функционального выражения тенденции ряда, а затем новых, расчетных значений ряда.

**Методы “механического” сглаживания.** К таким способам относятся следующие нижеприведенные.

**Метод усреднения по двум (нескольким) половинам ряда,** когда ряд делится на две (несколько) части. Затем, рассчитываются два (несколько) значения средних уровней ряда, по которым графически определяется тенденция ряда. Очевидно, что такой способ не достаточно полно отражает основную закономерность развития явления.

**Метод укрупнения интервалов,** при котором производится увеличение протяженности временных промежутков, и рассчитываются новые значения уровней ряда.

**Метод скользящей средней.** Данный метод применяется для характеристики тенденции развития исследуемой статистической совокупности и основан на расчете средних уровней ряда за определенный период. Последовательность определения скользящей средней такова:

1. Устанавливается интервал сглаживания или число входящих в него уровней. Если при расчете средней учитываются три уровня, скользящая средняя называется трехчленной, пять уровней – пятичленной и т.д. Если сглаживаются мелкие, беспорядочные колебания уровней в ряду динамики, то интервал (число скользящей средней) увеличивают. Если волны следует сохранить, число членов уменьшают.

2. Вычисляют первый средний уровень по арифметической простой:

$$y_n = \frac{y_1 + y_2 + \dots + y_n}{n},$$

где  $y_n$  – скользящая средняя,  $y_1, y_2, y_n$  – уровни ряда,  $n$  – членность скользящей средней.

3. Процесс продолжается до тех пор, пока в расчет  $y$  будет включен последний уровень исследуемого ряда динамики  $y_n$ .

4. По ряду динамики, построенному из средних уровней, выявляют общую тенденцию развития явления.

Отрицательной стороной использования метода скользящей средней является образование сдвигов в колебаниях уровней ряда, обусловленных “скольжением” интервалов укрупнения. Сглаживание с помощью скользящей средней может привести к появлению “обратных” колебаний, когда выпуклая “волна” заменяется на вогнутую.

В последнее время стала рассчитываться **адаптивная скользящая средняя**. Ее отличие состоит в том, что среднее значение признака, рассчитываемое так же, как описано выше, относится не к середине ряда, а к последнему промежутку времени в интервале укрупнения. Причем предполагается, что адаптивная средняя зависит от предыдущего уровня в меньшей степени, чем от текущего. То есть, чем больше промежутков времени между уровнем ряда и средним значением, тем меньшее влияние оказывает значение этого уровня ряда на величину средней.

**Метод экспоненциальной средней.** Экспоненциальная средняя – это адаптивная скользящая средняя, рассчитанная с применением весов, зависящих от степени “удаленности” отдельных уровней ряда от среднего значения. Величина веса

убывает по мере удаления уровня по хронологической прямой от среднего значения в соответствии с экспоненциальной функцией, поэтому такая средняя называется экспоненциальной. На практике применяется многократное экспоненциальное сглаживания ряда динамики, которое используется для прогнозирования развития явления.

**Методы “аналитического” выравнивания.** Более точным способом отображения тенденции динамического ряда является аналитическое выравнивание, т. е. выравнивание с помощью аналитических формул. В этом случае динамический ряд выражается в виде функции  $y(t)$ , в которой в качестве основного фактора принимается время  $t$ , и изменения аргумента функции определяют расчетные значения  $y_t$ . Фактическими (или эмпирическими) уровнями ряда динамики называют исходные данные об изменении явления, т. е. данные, полученные опытным путем, посредством наблюдения. Они обозначаются  $y_i$ . Расчетными (или теоретическими) уровнями ряда называют значения, полученные в результате подстановки в уравнение тренда значений  $t$ , и обозначают их.

Целью аналитического выравнивания динамического ряда является определение аналитической или графической зависимости  $f(t)$ . На практике по имеющемуся временному ряду задают вид и находят параметры функции  $f(t)$ , а затем анализируют поведение отклонений от тенденции. Функцию  $f(t)$  выбирают таким образом, чтобы она давала содержательное объяснение изучаемого процесса.

Чаще всего при выравнивании используются следующие зависимости:

- линейная  $f(t) = a_0 + a_1 t$
- параболическая  $f(t) = a_0 + a_1 t + a_2 t^2$
- экспоненциальная  $f(t) = \exp(a_0 + a_1 t)$
- сложная  $f(t) = \exp(a_0 + a_1 t + a_2 t^2)$

*Линейная зависимость* выбирается в тех случаях, когда в исходном временном ряду наблюдаются более или менее постоянные абсолютные и цепные приросты, не проявляющие тенденции ни к увеличению, ни к снижению.

*Параболическая зависимость* используется, если абсолютные цепные приросты сами по себе обнаруживают некоторую тенденцию развития, но абсолютные цепные приросты абсолютных цепных приростов (разности второго порядка) никакой тенденции развития не проявляют.

*Экспоненциальные зависимости* применяются, если в исходном временном ряду наблюдается либо более или менее постоянный относительный рост (устойчивость цепных темпов роста, темпов прироста, коэффициентов роста), либо, при отсутствии такого постоянства, - устойчивость в изменении показателей относительного роста (цепных темпов роста цепных же темпов роста, цепных коэффициентов роста цепных же коэффициентов или темпов роста и т.д.)

Графическое отображение изменения уровней ряда играет большую роль в применении данного вида выравнивания. Оно позволяет ускорить процедуру анализа и увеличить степень наглядности полученных результатов.

### **Задачи для самостоятельного решения**

**1.** Изобразите графически распределение месячного колебания вызовов скорой помощи детям при заболевании пневмонией. В январе зарегистрировано 100 вызовов, в феврале – 175, в марте – 153, в апреле – 138, в мае – 70, в июне – 51, в июле – 10, в августе – 21, в сентябре – 61, в октябре – 134, в ноябре – 191, в декабре – 136. В среднем за год – 100 вызовов скорой помощи.

**2.** Заболеваемость шигеллезом по месяцам года в абсолютных цифрах составляла: январь - 6; февраль - 9; март - 11; апрель - 10; май - 16; июнь - 23; июль - 19; август - 33; сентябрь - 58; октябрь - 19; ноябрь - 11; декабрь - 5. Всего за год - 220 случаев. Отобразите наглядно месячные колебания заболеваемости шигеллезом по сравнению со средним уровнем?

**3.** В Николаевском районе количество зарегистрированных больных язвой желудка за последние четыре года составляло: 1995 - 760 лиц, 1996 - 840 лиц, 1997 - 895 лиц, 1998 - 925 лиц. Среднегодовое количество населения и его структура за эти годы не изменялась. Охарактеризуйте тенденцию общей заболеваемости и ее темп в Николаевском районе.

### **Контрольные вопросы**

1. Применение динамических рядов в медицинской статистике
2. Динамический ряд. Уровни динамического ряда. Примеры
3. Классификация динамических рядов. Характеристика отдельных видов динамических рядов
4. Выравнивание динамического ряда. Цели. Задачи. Методы.

## **Рекомендованная литература**

1. Социальная гигиена и организация здравоохранения под редакцией А.Ф.Серенко и В.В.Ермакова.-Москва.-Медицина.-1984.-с.5-24, 102-113, 168-184.
2. Руководство к практическим занятиям по социальной гигиене и организации здравоохранения под редакцией Ю.П.Лисицина. - Москва. - Медицина.-1984.-с.49-52, 11-143.
3. Соціальна медицина та організація охорони здоров'я під редакцією Вороненка Ю.В., Москаленка В.Ф.- Тернопіль.-“Укрмедкнига”.-2000.
4. Соціальна медицина та організація охорони здоров'я під редакцією Голяченко О.М., Приходського О.О., Сердюка А.М.-Вінниця.- 1997.
5. Практикум по соціальній медицині та організації охорони здоров'я під редакцією Вороненка Ю.В.-Київ.-Здоров'я.-2001.
6. Збірник з тестових задач по соціальній медицині та організації охорони здоров'я під редакцією Вороненка Ю.В.-Київ.-Здоров'я.-1996.
7. Методичні рекомендації до практичних занять по соціальній медицині та організації охорони здоров'я.

## **Оглавление**

<b>ПРЕДИСЛОВИЕ .....</b>	<b>3</b>
<b>СОЦИАЛЬНАЯ МЕДИЦИНА И ОРГАНИЗАЦИЯ ЗДРАВООХРАНЕНИЯ. МЕДИЦИНСКАЯ СТАТИСТИКА. ОРГАНИЗАЦИЯ СТАТИСТИЧЕСКОГО ИССЛЕДОВАНИЯ .....</b>	<b>4</b>
<b>ОТНОСИТЕЛЬНЫЕ ВЕЛИЧИНЫ И ИХ ГРАФИЧЕСКОЕ ИЗОБРАЖЕНИЕ .....</b>	<b>28</b>
<b>ВАРИАЦИОННЫЕ РЯДЫ. СРЕДНИЕ ВЕЛИЧИНЫ. МЕТОДИКА ИХ ВЫЧИСЛЕНИЯ И ОЦЕНКИ.....</b>	<b>42</b>
<b>ОЦЕНКА ДОСТОВЕРНОСТИ ПОЛУЧЕННЫХ РЕЗУЛЬТАТОВ. МЕТОДЫ. ЗАДАЧИ. ОБЛАСТЬ ПРИМЕНЕНИЯ .....</b>	<b>52</b>
<b>МЕТОД СТАНДАРТИЗАЦИИ.....</b>	<b>66</b>
<b>КОРРЕЛЯЦИОННО – РЕГРЕССИОННЫЙ МЕТОД В ОЦЕНКЕ МЕДИКО- БИОЛОГИЧЕСКОЙ ИНФОРМАЦИИ .....</b>	<b>71</b>
<b>ДИНАМИЧЕСКИЕ РЯДЫ. АНАЛИЗ РЯДОВ ДИНАМИКИ. МЕДИКО- БИОЛОГИЧЕСКОЕ ЗНАЧЕНИЕ И ОБЛАСТЬ ПРИМЕНЕНИЯ .....</b>	<b>79</b>
<b>РЕКОМЕНДОВАННАЯ ЛИТЕРАТУРА .....</b>	<b>87</b>
<b>ОГЛАВЛЕНИЕ.....</b>	<b>88</b>